Skript zu den Vorlesungen Analysis I–III

Matthias Wendt

Inhaltsverzeichnis

Disclaimer: Wie üblich bei Vorlesungsmitschriften beruht nichts im folgenden Skript auf eigenen wissenschaftlichen Leistungen, Einsichten oder Ergebnissen. Sowohl Definitionen als auch Formulierungen und Beweise für Sätze sind meist das Produkt von ursprünglich vage formulierten Ideen, die durch vielfache Umformulierungen und Kontextveränderungen über Jahrzehnte zu den heute verwendeten Konzepten herangewachsen sind. Insofern können konkrete Zuschreibungen von Fragestellungen, Begriffen, Sätzen oder Beweistechniken üblicherweise nicht vorgenommen werden. (Außerdem stellt sich überraschend oft heraus, dass die üblicherweise mit Sätzen oder Begriffen verknüpften Namen nicht unbedingt die zentralen Entwicklungen des Begriffs reflektieren.)

Das Material des Vorlesungsskripts beruht überwiegend auf gängigen Lehrbüchern zur Analysis. Hauptsächlich verwendet wurden die Bücher von Königsberger [**koenigsberger**], Forster [**forster**] und Heuser [**heuser**]. Die einzige Eigenleistung beschränkt sich daher auf Tipp-, Index-, Logik- und andere Fehler, fragwürdige Anordnung des Materials sowie falsche Schwerpunktsetzungen. (Vermutlich sind allerdings viele meiner Fehler auch von anderen Leuten bereits gemacht worden; damit kann ich leben.)

KAPITEL 1

Mathematische Grundlagen

1.1. Aussagenlogik, Quantoren und Beweistechniken

Definition 1.1.1. *Eine* Aussage *ist ein sinnvolles sprachliches Gebilde, das wahr oder falsch sein kann.*

Beispiel 1.1.2.

- $1+1=2$ *(Aussage, wahr). Jede Quadratzahl ist durch 4 teilbar (Aussage, falsch).*
- *Draußen scheint die Sonne. (Aussage, Wahrheitswert situationsabhängig)*
- *Je zwei Geraden in der Ebene schneiden sich. (Aussage, falsch)*
- *n ist keine gerade Zahl. (keine Aussage, n nicht bestimmt, wird eine Aussage, wenn z.B. $n=4$ eingesetzt wird.) Ebenfalls sind Sätze mit* vielleicht, meistens *oder ähnlichen unpräzisen/unbestimmbaren Fomulierungen keine Aussagen.*
- *Diese Aussage ist nicht wahr. (keine Aussage, wegen Selbstreferenz, s. Russell-Paradox bzw. Lügner-Paradox)*

$\square$

Bemerkung 1.1.3. *Die Theorien in den empirischen Wissenschaften bestehen auch aus Aussagen. Aufgrund unseres begrenzten Zugangs zur Welt können wir aber die Wahrheit einer allgemeingültigen Aussage über die Welt (z.B. das Gravitationsgesetz von Newton) nie endgültig beweisen. Wir können nur immer wieder in unterschiedlichen Experimenten die Gültigkeit überprüfen. Deswegen ist das Theorie-Kriterium von Karl Popper auch Falsifizierbarkeit.*

Auch in der Mathematik gibt es Aussagen, von denen nicht bekannt ist, ob sie wahr oder falsch sind. Zum Beispiel war lange Zeit unbekannt, ob es für beliebige natürliche Zahlen $n \geq 3$ positive ganze Zahlen a, b, c gibt, die die Gleichung $a^n + b^n = c^n$ erfüllen.

Aussagen können logisch verknüpft werden:

"und": $A \wedge B$: Es gilt A und es gilt B.
"oder": $A \vee B$: Es gilt A oder es gilt B.
"nicht": $\neg A$: Es gilt nicht A (das Gegenteil von A).
Implikation: $A \Rightarrow B$: Aus A folgt B.
Äquivalenz: $A \Leftrightarrow B$: A gilt genau dann, wenn B gilt.

Beispiel 1.1.4.

(1) *Die Sonne scheint und es ist Nacht.*
(2) *Dieses Getränk ist Tee oder Kaffee (Achtung: inklusives vs exklusives Oder!)*
(3) *Wenn eine Zahl gerade ist, dann ist ihr Nachfolger ungerade.*
(4) *Jede Quadratzahl hat die Form $4n+1$.*

$\square$

Wahrheitswerte von komplexen Aussagen können mit Wahrheitstabellen bestimmt werden:

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
w	w	f	w	w	w	w
w	f	f	f	w	f	f
f	w	w	f	w	w	f
f	f	w	f	f	w	w

Bemerkung 1.1.5. *Hier heißen die A und B* Elementaraussagen

Übungsaufgabe 1.1.6. *Wie viele binäre logische Verknüpfungen gibt es? (binär=zweistellig, d.h. Verknüpfungen von zwei Aussagen)*

Zusammengesetzte Aussagen, die unabhängig vom Wahrheitswert der darin vorkommenden Elementaraussagen immer wahr sind, nennen wir *Tautologien.*

Übungsaufgabe 1.1.7. *Zeigen Sie mit Hilfe von Wahrheitstabellen, dass die folgenden Aussagen Tautologien sind:*

(1) $(\neg(A \wedge B)) \Leftrightarrow ((\neg A) \vee (\neg B))$.
(2) $(\neg(A \vee B)) \Leftrightarrow ((\neg A) \wedge (\neg B))$.
(3) $(A \Rightarrow B) \Leftrightarrow ((\neg A) \vee B)$.
(4) $(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$.
(5) $(A \wedge (B \vee C)) \Leftrightarrow ((A \wedge B) \vee (A \wedge C))$
(6) $((\neg A) \Rightarrow (B \wedge \neg B)) \Rightarrow A$
(7) $(\neg\neg A) \Leftrightarrow A$
(8) $(A \wedge (A \Rightarrow B)) \Rightarrow B$

Die nächste Stufe von Komplexität in mathematischen Aussagen erreichen wir durch Quantoren:

Existenzquantor: $\exists x : A(x)$ bedeutet "Es gibt ein x mit der Eigenschaft $A(x)$."
Allquantor: $\forall x : A(x)$ bedeutet "Für alle x gilt die Aussage $A(x)$."

Beispiel 1.1.8. *Die Reihenfolge der Quantoren spielt eine wesentliche Rolle, wie am Unterschied zwischen den beiden folgenden Aussagen zu erkennen ist:*

- *Für jede Person gibt es ein interessantes Studienfach.*
- *Es gibt ein Studienfach, das für alle Personen interessant ist.*

□

Zur Negation von Aussagen mit Quantoren:

- $\neg\exists x : A(x)$ ist gleichbedeutend mit $\forall x : \neg A(x)$.
- $\neg\forall x : A(x)$ ist gleichbedeutend mit $\exists x : \neg A(x)$.

Übungsaufgabe 1.1.9. *Was bedeutet die folgende Aussage? Beweisen Sie sie.*

$$\forall x \in \mathbb{N}\colon ((\exists y \in \mathbb{N}\colon x = y^2) \Rightarrow (\exists z \in \mathbb{N}\colon (x = 4z) \vee (x = 4z + 1))).$$

Dabei ist als Vorgriff $\mathbb{N}$ die Menge der natürlichen Zahlen.

Todo: mehr Übungsaufgaben

Übungsaufgabe 1.1.10. *Formulieren Sie die folgende Argumentation mit den Symbolen der Aussagenlogik: "Alle Katzen haben vier Beine. Ich habe vier Beine. Also bin ich eine Katze."*

Todo: Bemerkungen zu mathematischen Beweisen

1.2. Mengen, Relationen und Abbildungen

Definition 1.2.1.

- *Eine* Menge *ist eine Zusammenfassung bestimmter wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens zu einem Ganzen (G. Cantor, 1875).*
- *Die Objekte in einer Menge heißen* Elemente. *Notation: $x \in M$ bedeutet "x ist Element der Menge M", $x \notin M$ bedeutet "x ist nicht Element der Menge M".*
- *Zwei Mengen sind gleich, wenn sie die gleichen Elemente haben. Notation: $A = B$.*

Beispiel 1.2.2.

- *Die* leere Menge *wird mit $\emptyset$ bezeichnet.*
- *Es gibt zum Beispiel die Menge der Studierenden, die Menge der natürlichen Zahlen, die Menge der Atome im Universum, die Menge der Wirbeltiere oder die Menge der bekannten Arten von Wirbeltieren...*
- *Mengen können durch Aufzählung ihrer Elemente definiert werden: $\{1, 2, 3, 4\}$. Dabei kommt es auf Reihenfolge und Wiederholung nicht an, $\{1, 2, 3\} = \{3, 3, 3, 2, 2, 1\}$.*
- *Die Menge der natürlichen Zahlen ist $\mathbb{N} = \{0, 1, 2, 3, \dots\}$. Dabei müssen wir im nächsten Abschnitt genauer klären, was die drei Punkte bedeuten.*

□

Definition 1.2.3. *Eine Menge B ist eine* Teilmenge *von A, wenn alle Elemente von B in A enthalten sind. Notation $B \subset A$ oder $B \subseteq A$.*

Beispiel 1.2.4. *Mengen können durch Angaben von Eigenschaften beschrieben werden: $\{x \in M \mid A(x)\}$ bezeichnet die Menge der Elemente in M, die die Aussage $A(x)$ erfüllen. Wenn z.B. $\mathbb{N}$ die Menge der natürlichen Zahlen bezeichnet, dann ist $\{x \in \mathbb{N} \mid \exists n \in \mathbb{N} : x = n^2\}$ die Menge der Quadratzahlen. Das ist eine Teilmenge von $\mathbb{N}$.* □

Beispiel 1.2.5. *Die leere Menge ist in allen Mengen enthalten. Mit Quantoren formuliert ist $\emptyset \subseteq A$ die Aussage $\forall x \in \emptyset \colon x \in A$. Da es kein x mit $x \in \emptyset$ gibt, ist die Aussage für alle $x \in \emptyset$ erfüllt. Alternativ können wir formulieren, was das Gegenteil $\emptyset \not\subseteq A$ bedeutet: $\neg \forall x \in \emptyset \colon x \in A$ ist äquivalent zu $\exists x \in \emptyset \colon x \notin A$. Damit das wahr ist, müsste aber in $\emptyset$ ein Element existieren.* □

Aus gegebenen Mengen können mit Mengenoperationen neue Mengen konstruiert werden. Gegeben sei eine Menge M und zwei Teilmengen $A, B \subseteq M$.

Durchschnitt: $A \cap B := \{x \in M \mid x \in A \wedge x \in B\}$
Vereinigung: $A \cup B := \{x \in M \mid x \in A \vee x \in B\}$
Differenz: $A \setminus B := \{x \in M \mid x \in A \wedge x \notin B\}$
Komplement: $A^{\mathrm{c}} := \{x \in M \mid x \notin A\}$.

Bemerkung 1.2.6. *Hier ist Vorsicht geboten, da die Menge M in der Notation A^{c} nicht vorkommt. Es ist $A^{\mathrm{c}} = M \setminus A$. (Um z.B. zu bestimmen, was das Komplement der Menge aller Pinguine ist, müssen wir sagen, in welcher Obermenge wir die Pinguine betrachten – die Menge aller Tiere, die Menge der Vögel, oder der Lebewesen in der Antarktis...)*

Definition 1.2.7. *Kardinalität von endlichen Mengen Für eine endliche Menge M bezeichnet $\#M$ die Anzahl der Elemente von M.*

Wir zeigen eine einfache Aussage über Mengen.

Proposition 1.2.8. *Seien $A, B \subseteq M$ Teilmengen einer Menge M. Dann gilt $(A \cap B)^{\mathrm{c}} = A^{\mathrm{c}} \cup B^{\mathrm{c}}$.*

Beweis. Wir haben die folgenden äquivalenten Umformungen für $x \in M$:

$$\begin{aligned} x \in (A \cap B)^{\mathrm{c}} \quad &\Leftrightarrow \quad x \notin A \cap B \\ &\Leftrightarrow \quad \neg(x \in A \wedge x \in B) \\ &\Leftrightarrow \quad \neg(x \in A) \vee \neg(x \in B) \\ &\Leftrightarrow \quad x \in A^{\mathrm{c}} \vee x \in B^{\mathrm{c}}. \quad \square \end{aligned}$$

Bemerkung 1.2.9. *Für eine Gleichheit $X = Y$ von Mengen müssen wir die Inklusionen $X \subseteq Y$ und $Y \subseteq X$ beweisen. Hier können wir aber auch mit äquivalenten Umformungen arbeiten. Für eine Veranschaulichung der Situation mit einem Venn-Diagram, s. Abbildung 1.*

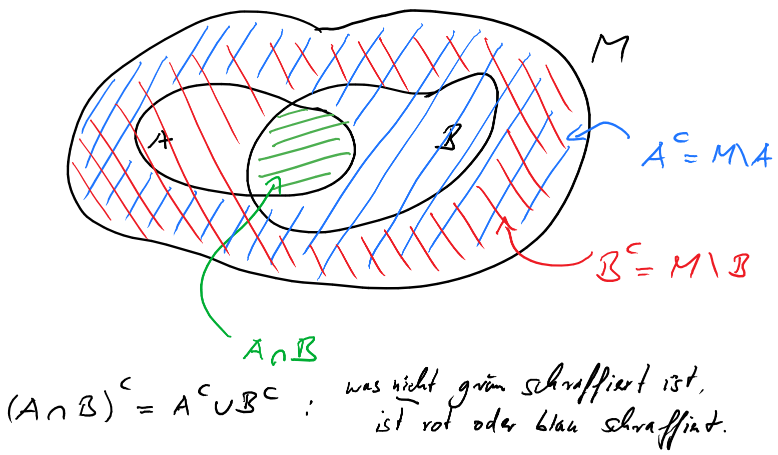


Abbildung 1. Venn-Diagramm für $(A \cap B)^{\mathrm{c}} = A^{\mathrm{c}} \cup B^{\mathrm{c}}$

Übungsaufgabe 1.2.10. *Gegeben seien drei Mengen A, B, C (alle Teilmengen einer Menge M). Sind die folgenden Aussagen richtig oder falsch? (Begründen Sie mit Beweis oder Gegenbeispiel)*

(a) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$
(b) $(A \setminus B) \cap C = (A \cap C) \setminus B$
(c) $(A \cup B) \cap A^{\mathrm{c}} = B$
(d) $A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$
(e) $A \setminus (B \cup C) = (A \setminus B) \cup (A \setminus C)$

Übungsaufgabe 1.2.11. *Für zwei Mengen A, B ist die symmetrische Differenz definiert durch $A \triangle B := (A \setminus B) \cup (B \setminus A)$. Zeigen Sie, dass die symmetrische Differenz ein Distributivgesetz bezüglich des Durchschnitts erfüllt, d.h. für drei Mengen A, B, C gilt*

$$A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C).$$

Definition 1.2.12. *Für Mengen A, B definieren wir die* Produktmenge *als Menge der geordneten Paare:*

$$A \times B := \{(a, b) \mid a \in A \wedge b \in B\}$$

Allgemeiner können wir für mehrere Mengen $A_1, \ldots, A_n$ eine Produktmenge definieren:

$$A_1 \times \cdots \times A_n := \{(a_1, \ldots, a_n) \mid a_i \in A_i \text{ für } i = 1, \ldots, n\}$$

Die Elemente $(a_1, \ldots, a_n)$ heißen geordnete n-Tupel.

DEFINITION 1.2.13. *Für Menge M definieren wir die Potenzmenge $\mathcal{P}(M)$, d.h. die Menge aller Teilmengen von M:*

$$\mathcal{P}(M) := \{A \mid A \subseteq M\}$$

PROPOSITION 1.2.14 (Russell-Paradox). *Sei R die Menge aller Mengen, die sich nicht selbst enthalten: $R = \{x \mid x \notin x\}$. Dann ist $R \in R$ genau dann, wenn $R \notin R$.*

BEWEIS. $\Rightarrow$: Da $R \in R$, ist R in sich selbst enthalten. Also kann R nicht in R enthalten sein.

$\Leftarrow$: Da $R \notin R$, ist R eine Menge, die sich nicht selbst enthält. Damit ist $R \in R$. □

BEMERKUNG 1.2.15. *Varianten: das Lügner-Paradox ist der Grund, warum "Dieser Satz ist falsch" keine Aussage ist.*

Bei der Konstruktion von Mengen müssen wir also vorsichtig sein, zum Beispiel kann es keine Menge von Mengen geben (sonst wäre R eine Teilmenge dieser Menge). Todo: allgemeine Bemerkung zu Zermelo–Fraenkel Mengenlehre

ÜBUNGSAUFGABE 1.2.16. *Sei V eine 4-elementige Menge. Seien $W_1, \ldots, W_4$ 2-elementige Teilmengen so dass für alle $\emptyset \neq I \subsetneq \{1, \ldots, 4\}$ gilt $\# \bigcap_{i \in I} W_i \leq 3 - \#I$. Bestimmen Sie*

$$\#\{U \subseteq V \mid \#U = 2, \forall i \in \{1, \ldots, 4\} : U \cap W_i \neq \emptyset\}$$

DEFINITION 1.2.17. *Seien A, B zwei Mengen. Eine* Abbildung *$f : A \to B$ ist eine Zuordnung, die jedem Element aus A genau ein Element aus B zuordnet:*

$$f : A \to B : a \mapsto f(a)$$

Dabei A der Definitionsbereich*, $f(a)$ der* Wert *der Abbildung für das* Argument *a.*

DEFINITION 1.2.18. *Der* Graph *einer Abbildung $f : A \to B$ ist definiert als $\Gamma(f) = \{(a, f(a)) \in A \times B \mid a \in A\} \subseteq A \times B$.*

BEMERKUNG 1.2.19. *Jede Teilmenge $\Gamma \subseteq A \times B$, für die es zu jedem $a \in A$ genau ein $b \in B$ mit $(a, b) \in \Gamma$ gibt, ist der Graph einer eindeutigen Abbildung.*

BEISPIEL 1.2.20.
- *Für jede Menge A haben wir die Identitätsabbildung $\mathrm{id}_A : A \to A : a \mapsto a$.*
- *$f : \mathbb{R} \to \mathbb{R} : x \mapsto x^2$.*
- *Die Reiseplanung für Bahn- Bus- oder Flugreisen kann als Abbildung aufgefasst werden, bei der jedem geordneten Tupel (S, Z) von Start- und Zielpunkt die Menge aller Verbindungen mit Start S und Ziel Z zugeordnet wird. Die Abbildungsvorschrift ist dabei im jeweiligen Auskunftssystem kodiert.*
- *Allgemein können wir uns Abbildungen gut prozedural vorstellen: eine Eingabe $a \in A$ wird verarbeitet, und das Ergebnis ist die Ausgabe $f(a) \in B$.*

Todo: mehr Beispiele □

DEFINITION 1.2.21. *Sei $f : A \to B$ eine Abbildung.*
- *Die Abbildung f heißt* injektiv*, wenn für alle $a_1, a_2 \in A$ aus $f(a_1) = f(a_2)$ schon $a_1 = a_2$ folgt.*
- *Die Abbildung f heißt* surjektiv*, wenn es für alle $b \in B$ ein $a \in A$ mit $f(a) = b$ gibt.*
- *Die Abbildung f heißt* bijektiv*, wenn sie injektiv und surjektiv ist.*

ÜBUNGSAUFGABE 1.2.22. *Untersuchen Sie die folgenden Abbildungen auf Injektivität, Surjektivität, Bijektivität:*

(1) $f_1\colon \mathbb{N} \to \mathbb{N}\colon n \mapsto 2n-1$

(2) $f_2\colon \mathbb{Z} \to \mathbb{Z}_{\leq 0}\colon n \mapsto \begin{cases} -n & n \geq 0 \\ n & n < 0 \end{cases}$

(3) $f_3\colon \mathbb{Z} \to \mathbb{N}\colon n \mapsto \begin{cases} 2n & n \geq 0 \\ -2n-1 & n < 0 \end{cases}$

(4) $f_4\colon \mathbb{Q} \to \mathbb{Q}_{\geq 0}\colon x \mapsto x^2$.

ÜBUNGSAUFGABE 1.2.23. *Seien $n, m \in \mathbb{N}$ zwei natürliche Zahlen. Zeigen Sie, dass es genau dann eine bijektive Abbildung $f\colon \{1, 2, \ldots, n\} \to \{1, 2, \ldots, m\}$ gibt, wenn $n = m$ ist.*

ÜBUNGSAUFGABE 1.2.24. *Seien X, Y und Z Mengen und seien $f\colon X \to Y$ und $g\colon Y \to Z$ Abbildungen.*

(1) Sei $g \circ f\colon X \to Z$ surjektiv. Ist g surjektiv? Ist f surjektiv?

(2) Sei $g \circ f\colon X \to Z$ injektiv. Ist g injektiv? Ist f injektiv?

Begründen Sie Ihre Antwort durch Beweis oder Gegenbeispiel.

DEFINITION 1.2.25. *Für eine Abbildung $f\colon A \to B$ und eine Teilmenge $C \subseteq A$ definieren wir*

$$f(C) = \{f(a) \mid a \in C\} \subseteq B.$$

Die Menge $f(A) \subseteq B$ wird auch als Bild *der Abbildung f bezeichnet, die Notation ist* $\mathrm{Im}(f) := f(A)$.

Für eine Abbildung $f\colon A \to B$ und eine Teilmenge $C \subseteq B$ definieren wir

$$f^{-1}(C) = \{a \in A \mid f(a) \in C\} \subseteq A.$$

Diese Menge heißt Urbildmenge *von C. Für $C = \{b\}$ mit $b \in B$ benutzen wir die Notation $f^{-1}(b) := f^{-1}(\{b\})$.*

ÜBUNGSAUFGABE 1.2.26. *Sei $f\colon M \to N$ eine Abbildung von Mengen und seien $A, B \subseteq M$ Teilmengen von M. Zeigen oder widerlegen Sie die folgenden Aussagen:*

(a) $f(A \cap B) = f(A) \cap f(B)$

(b) $f(A^{\mathrm{c}}) \subseteq f(A)^{\mathrm{c}}$

(c) $f(A^{\mathrm{c}}) \supseteq f(A)^{\mathrm{c}}$

ÜBUNGSAUFGABE 1.2.27. *Sei $f\colon X \to Y$ eine Abbildung. Zeigen Sie für Teilmengen $A, B \subseteq X$*

(1) Aus $A \subseteq B$ folgt $f(A) \subseteq f(B)$. (Gilt die Umkehrung?)

(2) $f(A \cap B) \subseteq f(A) \cap f(B)$.

(3) $f(A \cup B) = f(A) \cup f(B)$.

Zeigen Sie, dass $f(A \cap B) = f(A) \cap f(B)$ genau dann gilt, wenn f injektiv ist.

ÜBUNGSAUFGABE 1.2.28. *Sei $f\colon X \to Y$ eine Abbildung. Zeigen Sie für Teilmengen $A, B \subseteq Y$*

(1) Aus $A \subseteq B$ folgt $f^{-1}(A) \subseteq f^{-1}(B)$.

(2) $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$.

(3) $f^{-1}(A \cup B) = f(A)^{-1} \cup f^{-1}(B)$.

ÜBUNGSAUFGABE 1.2.29. *Sei $f\colon X \to Y$ eine Abbildung. Zeigen Sie für Teilmengen $A \subseteq X$ und $B \subseteq Y$*

(1) $f^{-1}(f(A)) \subseteq A$

(2) $f(f^{-1}(B)) \subseteq B$

ÜBUNGSAUFGABE 1.2.30. *Sei $f\colon A \to B$ eine Abbildung von Mengen. Wir definieren eine Abbildung $i\colon A \to \mathcal{P}(A)\colon a \mapsto \{a\}$. Zeigen Sie, dass das Bild der Abbildung*

$$f^{-1}\colon B \to \mathcal{P}(A)\colon b \in B \mapsto f^{-1}(b)$$

genau dann in $i(A) \subseteq \mathcal{P}(A)$ liegt, wenn f bijektiv ist.

BEMERKUNG 1.2.31. *Wenn f bijektiv ist, können wir die Abbildung $f^{-1}\colon B \to \mathcal{P}(A)$ als Abbildung $f^{-1}\colon B \to A$ auffassen. Diese Abbildung heißt* Umkehrabbildung *für f.*

DEFINITION 1.2.32. *Wir definieren die* Komposition *(das Zusammensetzen) von Abbildungen $f\colon M_1 \to M_2$ und $g\colon M_2 \to M_3$ durch $g \circ f(a) = g(f(a))$*

$$\begin{tikzcd} M_1 \arrow[r, "f"] \arrow[rr, bend right, "g\circ f"'] & M_2 \arrow[r, "g"] & M_3 \end{tikzcd}$$

ÜBUNGSAUFGABE 1.2.33. *Zeigen Sie, dass Komposition von Abbildungen assoziativ ist, d.h. $(h \circ g) \circ f = h \circ (g \circ f)$ für Abbildungen $f\colon A \to B$, $g\colon B \to C$ und $h\colon C \to D$.*

ÜBUNGSAUFGABE 1.2.34. *Sie $f\colon A \to B$ eine bijektive Abbildung. Zeigen Sie, dass genau eine Abbildung $g\colon B \to A$ existiert, so dass $g \circ f = \mathrm{id}_A$ und $f \circ g = \mathrm{id}_B$ gilt.*

1.3. Natürliche Zahlen und vollständige Induktion

Wir wollen die Menge $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ der natürlichen Zahlen definieren. Diese Menge ist unendlich, und müssen wir uns etwas mehr Gedanken machen, wie wir sie genau definieren.

DEFINITION 1.3.1 (Peano-Axiome). *(1) 0 ist eine natürliche Zahl.*
(2) Für jede natürliche Zahl n ist auch der Nachfolger $n+1$ eine natürliche Zahl.
(3) Zwei natürliche Zahlen sind genau dann gleich, wenn ihre Nachfolger gleich sind (d.h. die Nachfolgerabbildung $n \mapsto n+1$ ist injektiv).
(4) Es gibt keine natürliche Zahl n, deren Nachfolger 0 ist.
(5) Sei M eine Menge, die 0 enthält, und die für jede natürliche Zahl n auch deren Nachfolger $n+1$ enthält. Dann gilt $\mathbb{N} \subseteq M$.

Die so definierte Menge heißt die Menge der natürlichen Zahlen. Die Notation ist $\mathbb{N}$.

BEMERKUNG 1.3.2. *Ob 0 eine natürliche Zahl ist, ist einer der Standard-Streitpunkte in der Mathematik. Hier ist die 0 eine natürliche Zahl. Ich versuche, immer explizit zu sagen, ob 0 dabei ist oder nicht.*

ÜBUNGSAUFGABE 1.3.3. *Versuchen Sie, Beispiele für Mengen zu finden, die eins der Peano-Axiome verletzen, aber alle anderen erfüllen.*

Wenn wir nun genauer wissen, was die natürlichen Zahlen sind, stellt sich die Frage, wie wir Aussagen über natürliche Zahlen beweisen können. (Wenn wir eine Aussage $A(n)$ haben, in der eine natürliche Zahl n vorkommt, besteht $\forall n \in \mathbb{N}\colon A(n)$ eigentlich aus unendlich vielen Aussagen $A(0)$, $A(1)$, ..., die alle bewiesen werden müssen.) Die Antwort darauf gibt das

Prinzip der vollständigen Induktion:

Sei $A(n)$ eine Aussage, die von einer natürlichen Zahl n abhängt. Dann gilt die Aussage $A(n)$ für alle $n \in \mathbb{N}$, wenn
(1) $A(0)$ gilt, und
(2) für alle n die Implikation $A(n) \Rightarrow A(n+1)$ gilt.

BEMERKUNG 1.3.4. • *Die Aussage $A(0)$ nennen wir* Induktionsanfang. *In der Implikation $A(n) \Rightarrow A(n+1)$ heißt $A(n)$* Induktionsvoraussetzung *und $A(n+1)$* Induktionsbehauptung*, Implikation selbst ist der* Induktionsschritt*. Es ist immer gut, die Beweisstruktur klar zu formulieren.*

- *Induktion kann auch bei anderen natürlichen Zahlen anfangen.*
- *Ähnliche Beweisprinzipien gibt es auch für andere rekursiv definierten unendlichen Mengen, z.B. die Menge der Bäume.*

Ein einfaches Beispiel für einen Induktionsbeweis ist die folgende Aussage:

PROPOSITION 1.3.5. *Für alle natürlichen Zahlen ist die Summe der ersten n ungeraden Zahlen gleich n^2:*

$$1 + 3 + 5 + \cdots + (2n - 1) = n^2.$$

BEWEIS. Die Induktion beginnt hier bei $n = 1$, der Induktionsanfang ist mit $2 \cdot 1 - 1 = 1 = 1^2$ klar. Für den Induktionsschritt nehmen wir an, dass die Aussage für n gilt. Dann ist

$$1 + 3 + \cdots + (2n - 1) + (2n + 1) = n^2 + (2n + 1) = (n + 1)^2.$$

Im ersten Schritt haben wir die Induktionsvoraussetzung benutzt, im zweiten die binomische Formel. □

ÜBUNGSAUFGABE 1.3.6. *Zeigen Sie mit vollständiger Induktion, dass jede Quadratzahl von der Form $4n$ oder $4n + 1$ ist.*

Analog zum induktiven Beweis können wir auch durch natürliche Zahlen parametrisierte Objekte oder Notation definieren. Das nennen wir *rekursive Definition* (da sich die Definition des Begriffs auf den Begriff selbst rückbezieht): am Anfang wird das Objekt $X(0)$ definiert, und um das Objekt $X(n + 1)$ zu definieren, verwenden wir die Definition von $X(n)$.

BEISPIEL 1.3.7. *Potenzen mit natürlichen Exponenten können wir rekursiv definieren. Für eine (z.B. reelle) Zahl a definieren wir $a^0 := 1$ und der Rekursionsschritt definiert $a^{n+1} := a^n \cdot a$.* □

BEISPIEL 1.3.8. *Die Fakultät einer natürlichen Zahl n ist rekursiv definiert. Wir setzen $0! := 1$ und definieren $(n + 1)! = n! \cdot (n + 1)$.* □

BEISPIEL 1.3.9. *Notation für endliche Summen und Produkte kann ebenfalls rekursiv definiert werden.*

$$\begin{aligned} \sum_{i=m}^{n} a_i &= a_m + \cdots + a_n \\ \prod_{i=m}^{n} a_i &= a_m \cdot \cdots \cdot a_n \end{aligned}$$

□

ÜBUNGSAUFGABE 1.3.10. *Formulieren Sie die Definition von Summe und Produkt im Beispiel 1.3.9 als rekursive Definition.*

BEISPIEL 1.3.11. *Für eine etwas kompliziertere rekursive Definition definieren wir die Ordnung auf den natürlichen Zahlen ($m < n$, wenn m kleiner als n ist). Dafür definieren wir rekursiv die Mengen $\mathbb{N}_{<n}$, die genau die natürlichen Zahlen m enthalten, für die $m < n$ sein soll. Der Rekursionsanfang ist $\mathbb{N}_{<0} = \emptyset$, es gibt keine natürliche Zahl kleiner als 0. Im Rekursionsschritt setzen wir $\mathbb{N}_{<n+1} := \mathbb{N}_{<n} \cup \{n\}$.* □

ÜBUNGSAUFGABE 1.3.12. *Zeigen Sie, dass mit dieser Definition aus $n < m$ und $m < p$ schon $n < p$ folgt. Zeigen Sie, dass für alle Paare natürlicher Zahlen n, m mit $n \neq m$ entweder $n < m$ oder $m < n$ gilt.*

Satz 1.3.13 (Wohlordnungsprinzip der natürlichen Zahlen). *Jede nicht-leere Teilmenge der natürlichen Zahlen hat ein kleinstes Element.*

Beweis. Wir führen einen Widerspruchsbeweis. Sei dafür $\emptyset \neq M \subset \mathbb{N}$ eine Menge, die kein Minimum besitzt. Dann betrachten wir die Menge

$$U = \{k \in \mathbb{N} \mid k < m \text{ für alle } m \in M\}.$$

Wir zeigen durch vollständige Induktion, dass $U = \mathbb{N}$ ist.

Induktionsanfang $n = 0$**:** Wir zeigen $0 \in U$. Das Gegenteil $0 \notin U$ bedeutet, dass es $m \in M$ mit $0 \geq m$ gibt. Da 0 die kleinste natürliche Zahl ist, ist dann $m = 0$ und damit das kleinste Element der Menge M, was im Widerspruch zur Annahme steht. Also ist $0 \in U$.

Induktionsschritt: Sei $n \in U$, d.h. $n < m$ für alle $m \in M$. Wir wollen zeigen, dass $n + 1 \in U$ ist. Wenn $n + 1 \notin U$, dann existiert $m \in M$ mit $n + 1 \geq m$. Da $n < m$ für alle $m \in M$ und zwischen n und $n + 1$ keine weitere natürliche Zahl liegt, ist dann $n+1 = m$ und damit das Minimum von M. Wieder ein Widerspruch, also ist $n + 1 \in U$.

Nach dem Prinzip der vollständigen Induktion ist $U = \mathbb{N}$.

Nach Konstruktion ist $U \cap M = \emptyset$, und $U = \mathbb{N}$ widerspricht damit der Annahme $M \neq \emptyset$. Damit haben wir das Wohlordnungsprinzip bewiesen. □

Proposition 1.3.14. *Sei* $q \neq 1$ *eine reelle Zahl. Dann gilt für alle* $n \in \mathbb{N}$

$$\sum_{i=0}^{n} q^i = \frac{1 - q^{n+1}}{1 - q}.$$

Beweis. Wir beweisen die Aussage mit vollständiger Induktion über n. Für den Induktionsanfang $n = 0$ haben wir

$$\sum_{i=0}^{0} q^i = q^0 = 1 = \frac{1 - q^{0+1}}{1 - q}.$$

Im Induktionsschritt setzen wir die geometrische Summe für n voraus und rechnen

$$\begin{aligned}
\sum_{i=0}^{n+1} q^i &= \left(\sum_{i=0}^{n} q^i\right) + q^{n+1} \\
&= \frac{1 - q^{n+1}}{1 - q} + q^{n+1} \\
&= \frac{1 - q^{n+1} + (1 - q)q^{n+1}}{1 - q} = \frac{1 - q^{n+2}}{1 - q}.
\end{aligned}$$

Dabei benutzen wir im Übergang zur zweiten Zeile die Induktionsvoraussetzung. Nach dem Prinzip der vollständigen Induktion gilt die Formel für die geometrische Summe für alle $n \in \mathbb{N}$. □

Bemerkung 1.3.15. *Die geometrische Summe kommt überall vor. Ein Beispiel ist die Belohnung für die Erfindung des Schachspiels:*

https://de.wikipedia.org/wiki/Sissa_ibn_Dahir.

Weitere Beispiele und Übungsaufgaben für Induktion gibt es in der folgenden Auflistung. Es gibt auch einige Standard-Induktionsaufgaben im Umkreis der Fibonacci-Zahlen.

Beispiel 1.3.16. *Die Existenz einer Gewinnstrategie im Spiel "Nim" kann mit vollständiger Induktion über die Anzahl der Spielzüge gezeigt werden, siehe z.B.*

https://en.wikipedia.org/wiki/Nim

□

ÜBUNGSAUFGABE 1.3.17. *Sei $n > 0$ eine natürliche Zahl. Wir betrachten ein quadratisches Schachfeld im Format $2^n \times 2^n$, bei dem ein beliebiges Quadrat entfernt wird. Zeigen Sie, dass es möglich ist, den Rest des Schachfelds mit L-förmigen Kacheln von jeweils 3 Quadraten überschneidungsfrei zu kacheln.*

ÜBUNGSAUFGABE 1.3.18. *Zeigen Sie, dass die "Türme von Hanoi" mit n Scheiben in $2^n - 1$ Zügen lösbar sind. (Weitere Induktionsbeweise ergeben sich durch die Beziehung zum Sierpinski-Dreieck.)*

ÜBUNGSAUFGABE 1.3.19. *Bestimmen Sie die Anzahl der Tripel $(n_1, n_2, n_3) \in \mathbb{N}^3$ natürlicher Zahlen mit $n_1 + n_2 + n_3 = n + 1$ für $n \geq 2$.*

ÜBUNGSAUFGABE 1.3.20. *Zeigen Sie die folgenden Aussagen für alle natürlichen Zahlen $n \geq 1$:*

(1) $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$
(2) $\sum_{i=1}^{n} \frac{1}{i(i+1)} = \frac{n}{n+1}$
(3) $\prod_{i=1}^{n} \left(1 + \frac{1}{i}\right) = n + 1$

ÜBUNGSAUFGABE 1.3.21. *Zeigen Sie, dass es $n!$ Möglichkeiten gibt, die Zahlen $\{1, \dots, n\}$ anzuordnen.*

ÜBUNGSAUFGABE 1.3.22. *Zeigen Sie, dass die Anzahl der Elemente in $\mathcal{P}(\{1, \dots, n\})$ gleich 2^n ist.*

ÜBUNGSAUFGABE 1.3.23. *Zeigen Sie die folgenden Identitäten für Binomialkoeffizienten:*

$$\sum_{i=0}^{n} \binom{n}{i} = 2^n$$

$$\sum_{i=0}^{n} (-1)^i \binom{n}{i} = 0$$

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$$

ÜBUNGSAUFGABE 1.3.24. *Bestimmen Sie die Anzahl von Domino-Kachelungen eines $2 \times n$-Rechtecks.*

ÜBUNGSAUFGABE 1.3.25. *Sei $a_i = (4, 484, 48484, \dots)$ und $b_i = (8, 848, 84848, \dots)$.*

(1) *Zeigen Sie, dass für alle i gilt $4b_i - 7a_i = 4$.*
(2) *Sei c_i die Zusammensetzung von a_i und b_i, z.B. ist $c_1 = 48$ und $c_2 = 484848$. Zeigen Sie, dass für alle i gilt $c_i = b_i^2 - a_i^2$.*

ÜBUNGSAUFGABE 1.3.26.

- *Zeigen Sie, dass für alle natürlichen Zahlen n gilt $7 | (11^n - 4^n)$.*
- *Für welche n gilt $2^n \leq n!$.*
- *Zeigen Sie für jede natürliche Zahl $n \geq 1$, dass $9 \mid n^3 + (n+1)^3 + (n+2)^3$ gilt.*
- *Zeigen Sie, dass für alle natürlichen Zahlen $n \geq 3$ die Zahl $n^3 - 4n$ durch 3 teilbar ist.*

ÜBUNGSAUFGABE 1.3.27. *Zeigen Sie mit vollständiger Induktion für alle natürlichen Zahlen $n \geq 1$:*

$$(1 + 2 + 3 + \cdots + n)^2 = 1^3 + 2^3 + 3^3 + \cdots + n^3.$$

1.4. Reelle Zahlen: axiomatischer Ansatz und Körperaxiome

Grundlage der Analysis sind die reellen Zahlen. Für die Vorlesung werden wir die reellen Zahlen axiomatisch formulieren. Die Konstruktion der reellen Zahlen und der Nachweis, dass die Axiome auch tatsächlich erfüllt sind, werden in der Vorlesung "Grundlagen der Mathematik" durchgeführt. Auch für die Definition der ganzen Zahlen $\mathbb{Z}$ und der rationalen Zahlen $\mathbb{Q}$ verweisen wir hier nur auf die GdM-Vorlesung.

Bemerkung 1.4.1. Todo: Bemerkungen zu Axiomen und Axiomensystemen in der Mathematik

Definition 1.4.2. *Wir bezeichnen mit $\mathbb{R}$ einen vollständigen angeordneten Körper.*

Bemerkung 1.4.3. *Die Körperaxiome werden in Definition 1.4.4 formuliert, die Anordnungsaxiome in Definition 1.5.1 und das Vollständigkeitsaxiom in Definition 1.6.1. Die reellen Zahlen sind (in Logik zweiter Stufe) durch diese Axiome eindeutig charakterisiert.*

Es gibt andere Axiomatisierungen der reellen Zahlen, z.B. von Tarski:

https://en.wikipedia.org/wiki/Tarski%27s_axiomatization_of_the_reals

Die Körperaxiome beziehen sich auf die Grundrechenarten und deren Eigenschaften:

Definition 1.4.4. *Ein* Körper *K ist eine Menge mit zwei Operationen $+\colon K\times K\to K$ (Addition) und $\cdot\colon K\times K\to K$ (Multiplikation), so dass die folgenden Axiome erfüllt sind.*

- *Für die Addition haben wir die folgenden Axiome (die besagen, dass $(K,+)$ eine abelsche Gruppe ist):*
  - **Assoziativität:** *Für alle $a,b,c\in K$ gilt $(a+b)+c=a+(b+c)$.*
  - **neutrales Element:** *Es existiert ein Element $0\in K$, so dass für alle $a\in K$ gilt $a+0=0+a=a$.*
  - **inverse Elemente:** *Für jedes Element $a\in K$ existiert ein Element $(-a)\in K$, so dass gilt $a+(-a)=(-a)+a=0$.*
  - **Kommutativität:** *Für alle $a,b\in K$ gilt $a+b=b+a$.*
- *Für die Multiplikation haben wir die folgenden Axiome:*
  - **Assoziativität:** *Für alle $a,b,c\in K$ gilt $(a\cdot b)\cdot c=a\cdot(b\cdot c)$.*
  - **neutrales Element:** *Es existiert ein Element $1\in K$ mit $1\neq 0$, so dass für alle $a\in K$ gilt $a\cdot 1=1\cdot a=a$.*
  - **inverses Element:** *Für jedes Element $a\in K\setminus\{0\}$ existiert ein Element $a^{-1}\in K$, so dass gilt $a\cdot a^{-1}=a^{-1}\cdot a=1$.*
  - **Kommutativität:** *Für alle $a,b\in K$ gilt $a\cdot b=b\cdot a$.*
- *Die* ***Distributivität*** *beschreibt die Interaktion zwischen Addition und Multiplikation: für alle $a,b,c\in K$ gilt $a\cdot(b+c)=a\cdot b+a\cdot c$.*

Beispiel 1.4.5.
- *Beispiele für Körper sind die rationalen Zahlen $\mathbb{Q}$, die reellen Zahlen $\mathbb{R}$ und die komplexen Zahlen $\mathbb{C}$. Es gibt auch Körper, deren Elemente nicht einfach Zahlen sind, z.B. den Körper der rationalen Funktionen mit reellen Koeffizienten (den wir später noch sehen werden).*
- *Die ganzen Zahlen $\mathbb{Z}$ sind kein Körper. Fast alle Axiome sind erfüllt, aber nicht alle ganzen Zahlen haben Inverse bezüglich der Multiplikation.*

□

Übungsaufgabe 1.4.6. *Für eine Menge M hat die Potenzmenge $\mathcal{P}(M)$ zwei Operationen, die symmetrische Differenz $\triangle$ und den Durchschnitt $\cap$, die sich wie Addition bzw. Multiplikation verhalten.*

(1) *Zeigen Sie, dass bis auf die Existenz multiplikativer Inverser alle Körperaxiome erfüllt sind. (Eine solche Struktur heißt* kommutativer Ring.*)*
(2) *Zeigen Sie, dass für M eine ein-elementige Menge $\mathcal{P}(M)$ ein Körper ist. (Dieser Körper wird üblicherweise mit $\mathbb{F}_2$ bezeichnet.)*

Wir leiten ein paar elementare Konsequenzen aus den Körperaxiomen ab.

Lemma 1.4.7. (1) *Die neutralen Elemente 0 und 1 sind eindeutig bestimmt.*
(2) *Die additiven und multiplikativen Inversen sind eindeutig bestimmt.*
(3) *Eine lineare Gleichung $ax + b = c$ mit $a, b, c \in K$ und $a \neq 0$ hat eine eindeutige Lösung.*

Beweis. (1) Sei $0' \in K$ ein Element, so dass $x + 0' = x$ für alle $x \in K$ gilt. Dann haben wir $0 + 0' = 0$ (für den Fall $x = 0$). Da 0 ein neutrales Element ist, gilt $0 + 0' = 0' + 0 = 0'$. Zusammen haben wir $0 = 0 + 0' = 0'$. Die Eindeutigkeit von 1 wird analog gezeigt.

(2) Sei $y \in K$ ein Element mit $x + y = 0$. Addieren wir auf beiden Seiten der Gleichung $(-x)$ erhalten wir $x + y + (-x) = -x$. Die Eigenschaft der additiven Inversen und die Kommutativität liefern $0+y = x+(-x)+y = x+y+(-x) = -x$. Da 0 das neutrale Element ist, folgt $y = -x$. Die Eindeutigkeit der multiplikativen Inversen wird analog gezeigt.

(3) □

Bemerkung 1.4.8. *Die eindeutige Lösbarkeit linearer Gleichungen kodiert viele der relevanten Eigenschaften aus den Körperaxiomen und kann auch für alternative Axiomatisierung von Körpern benutzt werden.*

Übungsaufgabe 1.4.9 (Körperaxiome, Bruchrechnung). *Zeigen Sie die folgenden Regeln für das Bruchrechnen. Dabei sind a, b, c, d Elemente eines Körpers K, und es gilt $b \neq 0$ und $d \neq 0$. Die Notation ist wie in der Vorlesung $\frac{a}{b} := a \cdot b^{-1}$.*

(1) $\frac{a}{b} = \frac{c}{d}$ *genau dann, wenn $ad = bc$.*
(2) $\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}$
(3) $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$

Übungsaufgabe 1.4.10 (Konsequenzen der Körperaxiome). *Leiten Sie die folgenden Aussagen aus den Körperaxiomen ab. Hier ist K ein Körper.*

(1) *Für alle $a \in K$ ist $a \cdot 0 = 0$. (Diskutieren Sie, was das für Division durch 0 bedeutet.)*
(2) *Ein Körper ist* nullteilerfrei*, d.h. aus $xy = 0$ folgt $x = 0$ oder $y = 0$.*
(3) *Für alle $x \in K$ ist $-x = (-1) \cdot x$.*
(4) *Für alle $x, y \in K$ ist $(-x) \cdot (-y) = x \cdot y$*
(5) *Für alle $x \in K \setminus \{0\}$ ist $(x^{-1})^{-1} = x$.*
(6) *Für alle $x, y \in K$ mit $x, y \neq 0$ gilt $(x \cdot y)^{-1} = x^{-1} \cdot y^{-1}$.*

Übungsaufgabe 1.4.11. *Wo ist der Fehler im Beweis $2 = 1$?*

$$\begin{aligned} a = b \quad &\Rightarrow \quad a^2 = ab \\ &\Rightarrow \quad a^2 - b^2 = ab - b^2 \\ &\Rightarrow \quad (a+b)(a-b) = b(a-b) \\ &\Rightarrow \quad a + b = b \\ &\Rightarrow \quad 2b = b \\ &\Rightarrow \quad 2 = 1. \end{aligned}$$

Todo: mehr Übungsaufgaben

BEISPIEL 1.4.12. *Zu diesem Zeitpunkt kennen wir noch nicht viele reelle Zahlen. Die meisten Körperaxiome formulieren Eigenschaften von Addition und Multiplikation, mit denen wir aus gegebenen Elementen neue Element produzieren können. Die einzigen Elemente, deren Existenz explizit in den Körperaxiomen gefordert wird, sind* 0 *und* 1. *Damit können wir für eine natürliche Zahl n ein Element $n \in \mathbb{R}$ wie folgt definieren:*

$$n := \underbrace{1 + 1 + \cdots + 1}_{n\text{-}mal}.$$

Dadurch erhalten wir eine Abbildung $\mathbb{N} \to \mathbb{R}\colon n \mapsto n$. Diese Abbildung erweitern wir zu einer Abbildung $\mathbb{Z} \to \mathbb{R}$, indem wir eine negative ganze Zahl $-n$ auf $-n \in \mathbb{R}$ abbilden. Aus den Körperaxiomen folgt aber nicht, dass diese Abbildung injektiv ist (da es zum Beispiel Körper mit zwei Elementen gibt, in denen $1 + 1 = 0$ gilt). Besonders viele Zahlen erhalten wir so nicht. Für interessantere Zahlen brauchen wir also mehr Axiome. □

1.5. Reelle Zahlen: Anordnungsaxiome

DEFINITION 1.5.1. *Ein Körper K heißt* angeordnet*, wenn es eine Teilmenge $< \subset K \times K$ gibt, so dass die folgenden Axiome gelten:*

Trichotomie: *Für $a, b \in K$ gilt genau eine der drei Beziehungen:*

$$a < b, \quad a = b, \quad a > b$$

Transitivität: *Für alle $a, b, c \in K$ folgt aus $a < b$ und $b < c$ auch $a < c$.*

Monotonie: *Aus $a < b$ folgt für alle $c \in K$ auch $a + c < b + c$. Aus $a < b$ folgt für alle $c > 0$ auch $ac < bc$.*

BEMERKUNG 1.5.2. *Die übliche Notation für Paare in der Teilmenge $<$ ist: $a < b$ bedeutet $(a, b) \in <$. Wir schreiben $a > b$ für $b < a$. Außerdem definieren wir $a \leq b$ als $(x < y) \vee (x = y)$, analog für $a \geq b$.*

Angeordnete Körper können auch durch die Teilmenge der positiven Elemente charakterisiert werden, dies liefert eine alternative Axiomatisierung:

PROPOSITION 1.5.3. *Ein Körper K ist genau dann angeordnet, wenn es eine Teilmenge $K_+ \subset K$ von* positiven Elementen *gibt, so dass die folgenden Aussagen gelten:*

Trichotomie: *Für jedes Element $x \in K$ gilt eine der drei Aussagen:*

$$x > 0, \quad x = 0, \quad -x > 0$$

Abgeschlossen unter Addition: *für $x, y > 0$ gilt $x + y > 0$.*

Abgeschlossen unter Multiplikation: *für $x, y > 0$ gilt $x \cdot y > 0$.*

BEWEIS. Die Übersetzung zwischen den Anordnungsaxiomen und den alternativen Axiomen in der Proposition erfolgt durch: $x > y$ genau dann, wenn $x - y > 0$. (Unter den Anordnungsaxiomen folgt das aus der Monotonie.)

Wir zeigen zuerst, dass aus den Anordnungsaxiomen folgt, dass $x > y$ äquivalent zu $-y > -x$ ist. Nach Monotonie reicht es, die Äquivalenz von $x - y > 0$ und $(-y) - (-x) > 0$ zu zeigen, die Gleichheit $x - y = (-y) - (-x)$ folgt aber aus den Körperaxiomen.

"$\Rightarrow$": Wir nehmen an, dass K angeordnet ist und zeigen die alternativen Axiome. Die Trichotomie ist weitgehend klar, wenn wir die Äquivalenz von $x < 0$ und $-x > 0$ aus der Vorüberlegung benutzen. Dass die positiven Elemente unter Addition abgeschlossen sind, folgt aus Monotonie und Transitivität: $x > 0$ ist äquivalent zu $x + y > y$, und mit $y > 0$ folgt $x + y > 0$. Dass die positiven Elemente unter

Multiplikation abgeschlossen sind, folgt mit Monotonie und Körperaxiomen: aus $x > 0$ folgt $xy > 0 \cdot y = 0$.

"$\Leftarrow$": Wir nehmen nun an, dass die alternativen Axiome gelten und folgern, dass K angeordnet ist. Die alternative Formulierung der Trichotomie liefert $x - y > 0$, $x-y = 0$ oder $-(x-y) > 0$ als drei sich gegenseitig ausschließende Fälle. Im letzten Fall haben wir $-(x - y) = y - x > 0$, was äquivalent zu $x < y$ ist, damit ergibt sich die Trichotomie in den Anordnungsaxiomen. Für die Transitivität nehmen wir $x > y$ und $y > z$ an. Das bedeutet $x - y > 0$ und $y - z > 0$, und wir folgern $x - y + y - z = x - z > 0$ (also $x > z$), da positive Elemente unter Addition abgeschlossen sind. Für den additiven Teil der Monotonie erhalten wir aus $x-y > 0$ die Ungleichung $(x + z) - (y + z) > 0$. Für den multiplikativen Teil der Monotonie folgern wir aus $x - y > 0$ und $z > 0$ die Ungleichung $xz - yz = (x - y)z > 0$. □

Lemma 1.5.4. *Für einen angeordneten Körper haben wir die folgenden Aussagen für alle Elemente $a, b, c, d \in K$:*

(1) $a > b$ und $c < 0$ impliziert $ac < bc$.
(2) $a > b$ und $c > d$ impliziert $a + c > b + d$
(3) $x^2 > 0$ für alle $x \in K \setminus \{0\}$, insbesondere $1 > 0$.

Beweis. (1) Wir benutzen die Äquivalenz von $c < 0$ und $-c > 0$ aus dem Beweis von Proposition 1.5.3. Aus $a > b$ und $-c > 0$ folgern wir mit Monotonie $(-c)a > (-c)b$, was äquivalent zu $ac < bc$ ist.

(2) Aus den beiden Ungleichungen $a > b$ und $c > d$ folgern wir mit Monotonie die beiden Ungleichungen $a + c > b + c$ und $b + c > b + d$. Mit Transitivität folgt die Behauptung $a + c > b + d$.

(3) Für $x \neq 0$ haben wir nach Trichotomie nur zwei Möglichkeiten: entweder $x > 0$ oder $x < 0$. In beiden Fällen ist $x^2 > x \cdot 0 = 0$ mit der Monotonie bzw. Aussage (1) des Lemmas. Der Spezialfall ergibt sich aus $1 \neq 0$ und $1^2 = 1$. □

Übungsaufgabe 1.5.5 (Konsequenzen Anordnung). *Leiten Sie die folgenden Aussagen aus den Körper- und Anordnungsaxiomen ab. Hier ist K ein angeordneter Körper.*

(a) Für $x, y \in K$ mit $0 < x < y$ gilt $0 < y^{-1} < x^{-1}$.
(b) Seien $a, b \in K$ und $a < b$, sei $\lambda \in K$ mit $0 < \lambda < 1$. Dann gilt $a < \lambda a+(1-\lambda)b < b$. Jedes c mit $a < c < b$ ist von dieser Form.

Definition 1.5.6. *In einem angeordneten Körper K können wir für Elemente $x, y \in K$ Maximum und Minimum definieren:*

$$\max(x, y) := \begin{cases} x & x \geq y \\ y & \text{sonst} \end{cases}$$

$$\min(x, y) := \begin{cases} x & x \leq y \\ y & \text{sonst} \end{cases}$$

Rekursiv können wir diese Definition auf endliche Mengen reeller Zahlen ausdehnen.

Übungsaufgabe 1.5.7. *Zeigen Sie, dass jede endliche Menge reeller Zahlen ein kleinstes Element (Minimum) und ein größtes Element (Maximum) hat.*

Proposition 1.5.8 (Bernoulli-Ungleichung). *Für $x \in \mathbb{R}$ mit $x \geq -1$ und $n \in \mathbb{N}$ gilt*

$$(1 + x)^n \geq 1 + nx.$$

Beweis. Wir beweisen die Aussage durch vollständige Induktion über n. Für den Induktionsanfang $n = 0$ haben wir

$$(1 + x)^0 = 1 \geq 1 = 1 + 0 \cdot x.$$

Für den Induktionsschritt setzen wir voraus, dass $(1+x)^n \geq 1+nx$ für $x \geq -1$ gilt. Dann ist

$$(1+x)^{n+1} = (1+x)^n(1+x) \geq (1+nx)(1+x) = 1+(n+1)x+nx^2 \geq 1+(n+1)x.$$

Die erste Ungleichung folgt dabei aus der Induktionsvoraussetzung zusammen mit $1+x \geq 0$, die zweite Ungleichung aus $nx^2 \geq 0$. Nach dem Prinzip der vollständigen Induktion gilt die Behauptung. □

Definition 1.5.9. *Für einen angeordneten Körper definieren wir den Absolutbetrag*

$$|x| := \begin{cases} x & \text{für } x \geq 0 \\ -x & \text{für } x < 0 \end{cases}$$

Proposition 1.5.10. *Der Absolutbetrag hat die folgenden Eigenschaften:*

positive Definitheit: $|x| \geq 0$ *für alle* $x \in K$,
und $|x| = 0$ *genau dann, wenn* $x = 0$.
Multiplikativität: $|x \cdot y| = |x| \cdot |y|$ *für alle* $x, y \in K$.
Dreiecks-Ungleichung: $|x+y| \leq |x| + |y|$ *für alle* $x, y \in K$.

Beweis. Übungsaufgabe □

Wir diskutieren noch das Axiom des Archimedes. Dieses Axiom wird manchmal mit zu den Anordnungsaxiomen gezählt. Wir werden später sehen, dass es eine Konsequenz aus Vollständigkeit und Anordnung ist, d.h. dass ein vollständiger angeordneter Körper das Axiom des Archimedes erfüllt.

Definition 1.5.11. *Ein angeordneter Körper K heißt* archimedisch angeordnet*, wenn gilt:*

> ***Axiom des Archimedes:*** *Für je zwei Element $a, b \in K$ mit $a, b > 0$ gibt es eine natürliche Zahl n, so dass $na > b$.*

Dieses Axiom hat einige für die Analysis sehr relevante Konsequenzen.

Proposition 1.5.12. *Sei K ein archimedisch angeordneter Körper. Dann gelten die folgenden Aussagen:*

(1) Für jedes $\epsilon > 0$ existiert eine natürliche Zahl $n > 0$ mit $\frac{1}{n} < \epsilon$.
(2) Für $b > 1$ existiert für jedes Element $x \in K$ ein $n \in \mathbb{N}$, so dass $b^n > x$.
(3) Für $0 < b < 1$ existiert für jedes $\epsilon > 0$ ein $n \in \mathbb{N}$, so dass $b^n < \epsilon$.

Beweis. (1) ist eigentlich nur eine Umformulierung. Das Axiom besagt, dass für jedes Element $a \in K$ eine natürliche Zahl $n \in \mathbb{N}$ mit $n > a$ existiert. Für $\epsilon > 0$ haben wir zu $a = \frac{1}{\epsilon}$ eine natürliche Zahl $n > a$, damit ist dann $\frac{1}{n} < \frac{1}{a} = \epsilon$.

(2) Wir setzen $y = b-1 > 0$. Nach der Bernoulli-Ungleichung ist $b^n = (1+y)^n \geq 1+ny$. Aus dem Axiom des Archimedes folgt, dass ein $n \in \mathbb{N}$ mit $ny > x-1$, und wir erhalten $b^n > x$.

(3) erhalten wir aus (2) durch Kehrwerte bilden. □

Division mit Rest

Beispiel 1.5.13. *Nach der Diskussion der Anordnungsaxiome erhalten wir nun weitere Beispiele für reelle Zahlen (obwohl die Anordnungsaxiome keine Existenzaussagen beinhalten). Die zentrale Aussage hier ist $1 > 0$, aus der wir mit Monotonie und Transitivität (und einem induktiven Beweis) $n > 0$ für alle $n \in \mathbb{N}_{>0}$ folgern. Genauer können wir aus $1 > 0$ folgern, dass die Abbildung $\mathbb{Z} \to \mathbb{R}$ aus Beispiel 1.4.12 injektiv ist. Dann können wir diese Abbildung zu einer Abbildung $\mathbb{Q} \to \mathbb{R}$ fortsetzen, indem ein Bruch $\frac{p}{q}$ auf $p \cdot q^{-1} \in \mathbb{R}$ abgebildet wird. Mehr Beispiele für reelle Zahlen können wir auf diese Weise nicht erhalten, da die rationalen Zahlen auch ein angeordneter Körper sind.* □

ÜBUNGSAUFGABE 1.5.14. *Sei d eine ganze Zahl, die nicht durch eine Quadratzahl $p^2 > 1$ teilbar ist. Wir definieren auf $K := \mathbb{Q}^2$ eine Addition und Multiplikation:*

$$\begin{aligned}(a_1, b_1) + (a_2, b_2) &:= (a_1 + a_2, b_1 + b_2)\\ (a_1, b_1) \cdot (a_2, b_2) &:= (a_1 a_2 + d b_1 b_2, a_1 b_2 + a_2 b_1).\end{aligned}$$

Zeigen Sie, dass $(K, +, \cdot)$ ein Körper ist, der genau dann angeordnet ist, wenn d positiv ist. (Dabei dürfen Sie natürlich voraussetzen, dass $\mathbb{Q}$ ein Körper ist.)

1.6. Reelle Zahlen: Vollständigkeit

Wir diskutieren noch die Vollständigkeit der reellen Zahlen.

DEFINITION 1.6.1. *Ein angeordneter Körper heißt* vollständig *(besser: Dedekind-vollständig), wenn die folgende Eigenschaft erfüllt ist:*

> ***Supremumseigenschaft:*** *Jede nicht-leere nach oben beschränkte Menge reeller Zahlen besitzt eine kleinste obere Schranke.*

DEFINITION 1.6.2.
- *Eine Menge $M \subseteq \mathbb{R}$ heißt* nach oben beschränkt*, wenn ein $s \in \mathbb{R}$ existiert, so dass $m \leq s$ für alle $m \in M$ gilt. Dieses s heißt* obere Schranke*. Eine kleinste obere Schranke für M heißt* Supremum*. Die Notation ist* $\sup(M)$
- *Eine Menge $M \subseteq \mathbb{R}$ heißt* nach unten beschränkt*, wenn ein $s \in \mathbb{R}$ existiert, so dass $s \leq m$ für alle $m \in M$ gilt. Dieses s heißt* untere Schranke*. Eine größte untere Schranke heißt* Infimum*. Die Notation ist* $\inf(M)$.
- *Eine Menge $M \subseteq \mathbb{R}$ heißt* beschränkt*, wenn sie nach oben* und *unten beschränkt ist.*
- *Besitzt eine Menge $M \subseteq \mathbb{R}$ ein größtes Element, dann heißt dieses Element* Maximum*. Wenn ein Maximum existiert, ist diese Maximum dann auch das Supremum. Analog heißt ein kleinstes Element einer Menge $M \subseteq \mathbb{R}$ Minimum. Wenn ein Minimum existiert, ist dieses Element dann auch das Infimum.*

BEMERKUNG 1.6.3.
- *Für eine nach oben beschränkte Menge M mit oberer Schranke s ist jedes $s' \geq s$ auch eine obere Schranke, es gibt also unendlich viele obere Schranken. Suprema müssen dagegen nicht existieren. Wenn sie existieren, sind sie nach der Trichotomie der Anordnung eindeutig.*
- *Für eine nach oben beschränkte Menge $M \subseteq \mathbb{R}$ mit oberer Schranke s ist $s = \sup(M)$ genau dann, wenn für alle $\epsilon > 0$ ein $x \in M$ mit $s - \epsilon < x$ existiert.*

BEMERKUNG 1.6.4. *Andere Formulierungen der Vollständigkeit sind das* Intervallschachtelungsprinzip *(s. Satz 1.6.24) bzw. die Konvergenz von Cauchy-Folgen (im Kapitel zu Konvergenz).*

BEMERKUNG 1.6.5. *Es gibt auch eine Infimumseigenschaft für nach unten beschränkte Mengen. Diese ist wegen $\inf(-M) = -\sup(M)$ äquivalent zur Supremumseigenschaft.*

ÜBUNGSAUFGABE 1.6.6. *Seien $A, B \subset \mathbb{R}$ nach oben beschränkte Mengen. Zeigen Sie die folgenden Aussagen:*

(a) $\sup(A) + \sup(B) = \sup\{a + b \mid a \in A, b \in B\}$
(b) $-\sup(A) = \inf\{-a \mid a \in A\}$

ÜBUNGSAUFGABE 1.6.7. *Seien $A, B \subset \mathbb{R}$ nach oben beschränkte Mengen. Zeigen Sie, dass $\sup(A) \cdot \sup(B) = \sup\{a \cdot b \mid a \in A, b \in B\}$, falls alle Elemente von A und B nichtnegativ sind. Zeigen Sie, dass die Zusatzvoraussetzung notwendig ist.*

ÜBUNGSAUFGABE 1.6.8. *Für eine nach unten beschränkte Menge mit* $\inf(A) > 0$ *gilt* $(\inf(A))^{-1} = \sup\{a^{-1} \mid a \in A\}$.

ÜBUNGSAUFGABE 1.6.9. *Seien* $A, B \subseteq \mathbb{R}$ *nichtleere nach oben beschränkte Mengen. Zeigen Sie, dass* $A \cup B$ *nach oben beschränkt ist, und dass*

$$\sup(A \cup B) = \max\{\sup A, \sup B\}.$$

ÜBUNGSAUFGABE 1.6.10. *Bestimmen Sie (falls vorhanden) Infimum, Supremum, Minimum und Maximum der folgenden Teilmengen von reellen Zahlen:*
(a) $\left\{\frac{1}{m} + \frac{1}{n} \mid m, n \in \mathbb{N}_{>0}\right\}$.
(b) $\left\{x + \frac{1}{x} \mid \frac{1}{2} < x \leq 2\right\}$. *(Was ändert sich wenn wir stattdessen* $\frac{1}{2} < x < 2$ *fordern?)*

ÜBUNGSAUFGABE 1.6.11. *Untersuchen Sie, ob die folgenden Mengen ein Infimum, Supremum, Minimum bzw. Maximum haben und bestimmen Sie gegebenenfalls diese Werte:*
(a) $A = \{\frac{1}{x} - \frac{1}{y} \mid x, y \in \mathbb{R}, x, y \geq 1\}$
(b) $B = \{\frac{|x|x^2}{1+x^2} \mid x \in \mathbb{R}\}$
(c) $C = \{\frac{1}{1+n} + \frac{1+(-1)^n}{2n} \mid n \in \mathbb{N}, n \geq 1\}$
(d) $D = \{x \in \mathbb{R} \mid \exists y \in \mathbb{R} \colon (x+2)^2 + 4y^2 < 9\}$

PROPOSITION 1.6.12. *In einem vollständigen angeordneten Körper* K *gilt das Axiom des Archimedes, d.h. zu jedem* $x \in K$ *existiert* $n \in \mathbb{N}$ *mit* $x < n$.

BEWEIS. Wir nehmen an, dass das Axiom nicht erfüllt ist. Dann existiert $x \in K$ so dass für alle $n \in \mathbb{N}$ gilt $n < x$, d.h. $\mathbb{N}$ ist beschränkt. Aus der Vollständigkeit (Supremumseigenschaft) folgt die Existenz eines Supremums $s = \sup(\mathbb{N})$. Dann ist $s-1$ keine obere Schranke für $\mathbb{N}$, also existiert $n \in \mathbb{N}$ mit $s-1 < n$ bzw. $n+1 > s$. Da s eine obere Schranke für $\mathbb{N}$ war, haben wir einen Widerspruch, also gilt das Axiom des Archimedes. $\square$

BEMERKUNG 1.6.13. Todo: Bemerkungen zu Widerspruchsbeweis

ÜBUNGSAUFGABE 1.6.14 (Dichtheit der rationalen Zahlen). *Zeigen Sie, dass für jede reelle Zahl* ρ *und jedes* $\epsilon > 0$ *eine rationale Zahl* r *mit* $\rho - \epsilon < r < \rho + \epsilon$ *existiert. (Hinweis: Verwenden Sie das archimedische Axiom und das Wohlordnungsprinzip.)*

BEISPIEL 1.6.15. *Mit dem Vollständigkeitsaxiom erhalten wir nun gleich mehrere Möglichkeiten, reelle Zahlen zu definieren:*

durch Supremumseigenschaft: $\sqrt{2} = \sup\{x \in \mathbb{Q} \mid x^2 < 2\}$

durch Intervallschachtelungen: *Für die Kreiszahl* π *(das Verhältnis von Umfang zu Durchmesser eines Kreises) existieren Intervallschachtelungen, die auf Näherungen über Fläche (Archimedes) oder Umfang (Cusanus) von ein- bzw. umbeschriebenen n-Ecken basieren.*

durch konvergente Folgen/Reihen: *Eulersche Zahl*

$$\mathrm{e} = \sum_{i=0}^{\infty} \frac{1}{i!} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$$

$\square$

Als Beispiel für eine Definition reeller Zahlen durch Suprema haben wir die folgende Aussage zur Existenz von Quadratwurzeln positiver Zahlen.

PROPOSITION 1.6.16. *Für* $d \in \mathbb{R}$ *mit* $d > 0$ *betrachten wir die Menge* $S = \{x \in \mathbb{R} \mid x^2 \leq d\}$. *Für das Supremum* $s = \sup(S)$ *gilt* $s^2 = d$.

Beweis. Zuerst sehen wir, dass die Menge S nichtleer ($0 \in S$) und durch $\max(1, d)$ nach oben beschränkt ist. Nach der Vollständigkeit von $\mathbb{R}$ (in Form der Supremumseigenschaft) existiert das Supremum. Es bleibt zu zeigen, dass $s^2 = d$ ist.

Falls $s^2 > d$ ist, setzen wir $\epsilon = \frac{s^2-d}{2s} > 0$. Dann ist

$$(s-\epsilon)^2 = s^2 - 2s\epsilon + \epsilon^2 = s^2 - (s^2-d) + \epsilon^2 = s^2 - s^2 + d + \epsilon > d.$$

Damit ist $s - \epsilon$ eine obere Schranke, im Widerspruch zu $s = \sup(S)$.

Falls $s^2 < d$ ist, setzen wir $\delta = d - s^2 > 0$. Wir wählen ein $\epsilon > 0$ mit $\epsilon < 2s$ und $\epsilon < \frac{\delta}{4s}$. Das ist möglich mit dem archimedischen Axiom (da es natürliche Zahlen gibt, die größer als $\frac{1}{2s}$ und $\frac{4s}{\delta}$ sind). Mit dieser Wahl ist dann

$$(s+\epsilon)^2 = s^2 + 2s\epsilon + \epsilon^2 < s^2 + 2s\epsilon + 2s\epsilon < s^2 + 4s\frac{\delta}{4s} = s^2 + d - s^2 = d$$

Die erste Ungleichung folgt aus $\epsilon < 2s$, die zweite aus $\epsilon < \frac{\delta}{4s}$. Also ist $s + \epsilon \in S$, das widerspricht der Voraussetzung, dass s obere Schranke für S ist.

Aus der Trichotomie (Anordnungsaxiome für $\mathbb{R}$) folgt dann $s^2 = d$, was wir beweisen wollten. □

Bemerkung 1.6.17. *Die Menge $\{x \in \mathbb{Q} \mid x^2 < d\}$ hat das gleiche Supremum wie die Menge in Proposition 1.6.16. Dafür muss der obige Beweis abgewandelt werden, indem rationalen Zahlen statt $s \pm \epsilon$ für die Abschätzungen genutzt werden. Dies ist möglich, da die rationalen Zahlen in den reellen Zahlen dicht liegen, s. Übungsaufgabe 1.6.14.*

Proposition 1.6.18. *Die Zahl $\sqrt{2}$ ist irrational.*

Beweis. Wir beweisen die Behauptung durch Widerspruch. Nehmen wir an, dass $\sqrt{2}$ rational ist, d.h. dass es natürliche Zahlen p und $q \neq 0$ gibt, so dass $\frac{p}{q} = \sqrt{2}$ (oder äquivalent: $p^2 = 2q^2$) gilt.

Wenden wir das Wohlordnungsprinzip 1.3.13 auf die Menge $\{p \in \mathbb{N} \mid \exists q \in \mathbb{N} : p^2 = 2q^2\}$ an, sehen wir, dass ein kleinstes p existieren muss, für das es eine Darstellung $\frac{p}{q} = \sqrt{2}$ gibt.[1]

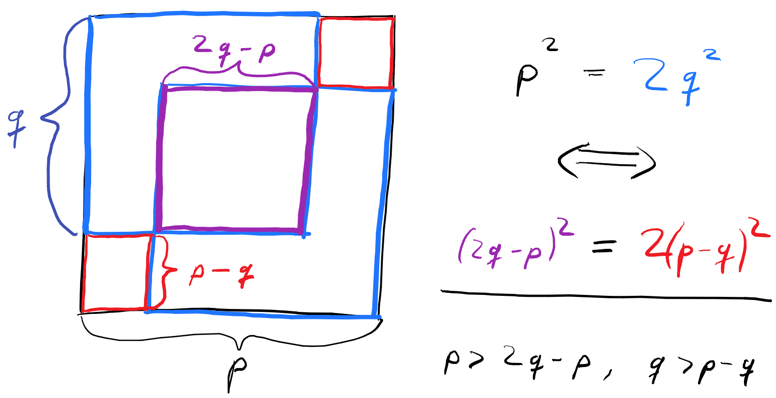


Abbildung 2. Bierdeckelbeweis für Irrationalität von $\sqrt{2}$

Aus der Abbildung 2 sehen wir aber, dass auch $(2q-p)^2 = 2(p-q)^2$ gilt. Aus $p^2 = 2q^2$ folgt $p > q$: wäre $p \leq q$, dann folgt mit $q \neq 0$ und $2 > 1$ schon $p^2 \leq q^2 < 2q^2$, was $p^2 = 2q^2$ widerspricht. Aus $p > q$ folgt dann $p > 2q - p$, also

[1]Das ist dann der gekürzte Bruch, das ist aber für den Beweis nicht wichtig.

finden wir für jede Darstellung von $\sqrt{2}$ als rationale Zahl $\frac{p}{q}$ eine neue Darstellung $\frac{2q-p}{p-q}$ mit kleinerem Zähler.

Damit haben wir einen Widerspruch, also ist $\sqrt{2}$ irrational. □

Bemerkung 1.6.19. *Zur Interpretation der Abbildung 2: p^2 ist die Fläche des großen Quadrates, q^2 die Fläche eines blauen Quadrates. Wenn die beiden blauen Quadrate den gleichen Flächeninhalt wie das große Quadrat haben sollen, d.h. $p^2 = 2q^2$, dann zählen wir das mittlere Quadrat doppelt, und lassen die beiden kleinen roten Quadrate aus. Damit sich diese Fehler in der Zählung wegheben, muss also $(2q-p)^2$, die Fläche des mittleren Quadrates, gleich $2(p-q)^2$, der Fläche der beiden kleinen Quadrate sein. Die Äquivalenz der beiden Gleichungen kann man auch rein algebraisch durch Ausmultiplizieren sehen, aber das Bild sagt halt mehr als tausend Worte.*

Definition 1.6.20 (Notation für Intervalle). *Für $a, b \in \mathbb{R}$ mit $a < b$ definieren wir die folgenden Intervalle:*

abgeschlossenes Intervall: $[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$
offenes Intervall: $(a, b) := \{x \in \mathbb{R} \mid a < x < b\}$
halboffene Intervalle: $[a, b) := \{x \in \mathbb{R} \mid a \leq x < b\}$ *sowie* $(a, b] := \{x \in \mathbb{R} \mid a < x \leq b\}$

Für diese Intervalle I definieren wir die Länge durch $|I| = b - a$.

Es gibt ein paar Sonderfälle, die sogenannten uneigentlichen Intervalle*:*

$$(-\infty, a) := \{x \in \mathbb{R} \mid x < a\},$$

sowie mit analogen Definitionen $(-\infty, a]$, $[a, +\infty)$, $(a, +\infty)$ und $(-\infty, +\infty) = \mathbb{R}$.

Bemerkung 1.6.21. *Besonders häufig benutzte Intervalle bzw. Mengen reeller Zahlen sind die* ϵ-Umgebungen einer reellen Zahl a

$$\{x \in \mathbb{R} \mid |x - a| < \epsilon\} = (a - \epsilon, a + \epsilon)$$

Definition 1.6.22 (Intervallschachtelung). *Eine* Intervallschachtelung (I_n) *ist eine Folge $I_1, I_2, I_3, \ldots$ von abgeschlossenen Intervallen, so dass*

(1) $I_{n+1} \subseteq I_n$ für alle $n \in \mathbb{N}$, $n > 0$.
(2) Für jedes $\epsilon > 0$ existiert ein Intervall I_n mit Länge $|I_n| < \epsilon$.

Intervallschachtelungsprinzip: Zu jeder Intervallschachtelung in $\mathbb{R}$ existiert eine reelle Zahl, die in allen Intervallen liegt.

Bemerkung 1.6.23. *Mit einem indirekten Beweis sehen wir, dass diese Zahl eindeutig bestimmt ist: liegen nämlich zwei verschiedene Zahlen $\alpha < \beta$ in allen Intervallen, dann ist $[\alpha, \beta] \subseteq I_n$ für alle $n \in \mathbb{N}$ und damit $|I_n| \geq \beta - \alpha$. Dies ist ein Widerspruch zu Eigenschaft (2) in der Definition der Intervallschachtelung.*

Wir erhalten nun die folgende alternative Charakterisierung der Vollständigkeit der reellen Zahlen:

Satz 1.6.24. *Die Supremumseigenschaft ist äquivalent zum Intervallschachtelungsprinzip.*

Beweis. "$\Rightarrow$" Wir beweisen das Intervallschachtelungsprinzip aus der Supremumseigenschaft. Sei dafür (I_n) mit $I_n = [a_n, b_n]$ eine Intervallschachtelung. Die Menge $A = \{a_1, a_2, \ldots\}$ der unteren Intervallgrenzen ist dann beschränkt, da alle b_n obere Schranken sind. Für das Supremum $s = \sup(A)$ gelten für alle $n \in \mathbb{N}$ die Ungleichungen $a_n \leq s \leq b_n$. Also ist $s \in [a_n, b_n]$ für alle $n \in \mathbb{N}$, und wir haben das Intervallschachtelungsprinzip gezeigt.

"$\Leftarrow$" Wir beweisen die Supremumseigenschaft aus dem Intervallschachtelungsprinzip. Sei dafür A eine nach oben beschränkte Menge. Wir konstruieren eine Intervallschachtelung $I_n = [a_n, b_n]$, die $\sup(A)$ definiert. Das bedeutet, dass alle b_n obere Schranken für A sind, aber keines der a_n eine obere Schranke für A ist.

Die Intervallschachtelung wird rekursiv konstruiert. Der Rekursionsanfang ist $[a_1, b_1]$, indem b_1 eine beliebige obere Schranke ist, und a_1 irgendein Element, das keine obere Schranke ist, z.B. $x - 1$ für $x \in A$. Für den Rekursionsschritt halbieren wir das gegebene Intervall $[a_n, b_n]$. Der Mittelpunkt ist $m = \frac{a_n+b_n}{2}$, und wir setzen

$$[a_{n+1}, b_{n+1}] := \begin{cases} [a_n, m] & m \text{ obere Schranke für } A \\ [m, b_n] & m \text{ keine obere Schranke für } A \end{cases}$$

Dies ist eine Intervallschachtelung. Die Inklusionen in Bedingung (1) sind klar. Da die Intervalle in jedem Schritt halbiert werden, haben wir (mit einem kurzen Induktionsbeweis) $|I_n| = \left(\frac{1}{2}\right)^{n-1} |I_1|$. Für ein beliebiges $\epsilon > 0$ existiert dann nach dem archimedischen Axiom ein $n \in \mathbb{N}$ mit

$$\left(\frac{1}{2}\right)^{n-1} < \frac{\epsilon}{|I_1|}$$

und damit gilt $|I_n| < \epsilon$. Dies zeigt Bedingung (2) für die Intervallschachtelung. Außerdem sind nach Konstruktion alle b_n obere Schranken für A sind, aber keines der a_n eine obere Schranke für A. Wir bezeichnen mit s die in allen Intervallen $[a_n, b_n]$ liegende Zahl, die nach dem Intervallschachtelungsprinzip existiert.

Wir zeigen zuerst, dass s eine obere Schranke ist. Wenn s keine obere Schranke ist, dann existiert $x \in A$ mit $x > s$ und ein Intervall $[a_n, b_n]$ mit $b_n - a_n < x - s$. Wegen $s \in [a_n, b_n]$ folgt $b_n - s < x - s$, also $b_n < x$. Aber b_n ist nach Konstruktion eine obere Schranke für A, und wir erhalten einen Widerspruch. Also ist s eine obere Schranke.

Wir zeigen noch, dass s die kleinste obere Schranke ist. Wir nehmen wieder an, dass s nicht die kleinste obere Schranke ist, also ein $s' < s$ existiert, das auch eine obere Schranke ist. Dann existiert $[a_n, b_n]$ mit $b_n - a_n < s - s'$ und wegen $s \in [a_n, b_n]$ folgt $s - a_n < s - s'$, also $a_n > s'$. Da nach Konstruktion a_n keine obere Schranke ist, erhalten wir einen Widerspruch. Damit haben wir gezeigt, dass s ein Supremum für A ist. □

Wir können jetzt aus dem Intervallschachtelungsprinzip die Existenz von beliebigen Wurzeln ableiten:

Satz 1.6.25 (Existenz von k-ten Wurzeln). *Zu jeder reellen Zahl $x > 0$ und jeder natürlichen Zahl $k > 0$ existiert genau eine reelle Zahl $y > 0$ mit $y^k = x$.*

Bemerkung 1.6.26. *Die Zahl y aus dem Satz wird mit $y = \sqrt[k]{x}$ bezeichnet und k-te Wurzel genannt.*

Beweis. Es genügt, den Fall $x > 1$ zu zeigen. Für $x < 1$ ist die k-te Wurzel der Kehrwert der k-ten Wurzel aus $x^{-1} > 1$.

Wir konstruieren eine Intervallschachtelung $I_n = [a_n, b_n]$, so dass für alle $n > 0$ gilt $a_n^k \leq x \leq b_n^k$. Die Intervallschachtelung wird rekursiv konstruiert. Der Rekursionsanfang ist $I_1 = [1, x]$. Wegen $1 < x$ (und damit $x \leq x^k$ für alle natürlichen $k \geq 1$) ist $a_1^k \leq x \leq b_1^k$ erfüllt. Im Rekursionsschritt halbieren wir das gegebene Intervall $I_n = [a_n, b_n]$. Der Mittelpunkt ist $m = \frac{a_n+b_n}{2}$ und wir setzen

$$[a_{n+1}, b_{n+1}] := \begin{cases} [a_n, m] & m^k \geq x \\ [m, b_n] & m^k < x \end{cases}$$

Für die so konstruierten Intervalle gilt $a_n^k \leq x \leq b_n^k$ für alle $n > 0$. Wir erhalten eine Intervallschachtelung, da die Intervalllängen immer halbiert werden (Argument wie im Beweis von Satz 1.6.24). Also existiert nach dem Intervallschachtelungsprinzip eine Zahl y, die in allen Intervallen liegt.

Um zu zeigen, dass $y^k = x$ betrachten wir eine weitere Intervallschachtelung $I_n^k = [a_n^k, b_n^k]$. Bedingung (1) für Intervallschachtelung ist $I_{n+1}^k \subseteq I_n^k$, was aus $I_{n+1} \subseteq I_n$ folgt. Für Bedingung (2) haben wir

$$|I_n^k| = b_n^k - a_n^k = (b_n - a_n)(b_n^{k-1} + b_n^{k-2}a_n + \cdots + a_n^{k-1}).$$

Das zweite Gleichheitszeichen folgt mit einem Induktionsbeweis ähnlich wie der Beweis der geometrischen Summe. Damit können wir (unter Benutzung von $a_n < b_n$ und $b_1 \geq b_n$) abschätzen $|I_n^k| < |I_n| \cdot kb_1^{k-1}$. Da (I_n) eine Intervallschachtelung war, haben wir für ein gegebenes $\epsilon > 0$ einen Index ν s.d. $|I_\nu| < \frac{\epsilon}{kb_1^{k-1}}$ und damit $|I_\nu^k| < \epsilon$. Also gilt Bedingung (2), und (I_n^k) ist eine Intervallschachtelung. Für diese Intervallschachtelung liegen jetzt sowohl x (wegen $a_n^k \leq x \leq b_n^k$) als auch y^k (wegen $y \in I_n$) in I_n^k für alle n. Damit folgt $x = y^k$.

Zuletzt müssen wir noch die Eindeutigkeit von y zeigen. Für ein $z > 0$ mit $z^k = x$ folgt aus $z > y$ auch $z^k > y^k$, was im Widerspruch zu $z^k = x = y^k$ steht.

Damit haben wir die Existenz von eindeutigen Wurzeln gezeigt. □

Bemerkung 1.6.27. Todo: Bemerkung Beweisstruktur
zur Übung/zum Verständnis: Übersetzung in Existenz von Wurzeln mit Supremumseigenschaft

Bemerkung 1.6.28. Todo: Bemerkung zu Dezimalbruchentwicklung und berechenbaren reellen Zahlen

Übungsaufgabe 1.6.29. *(arithmetisches vs geometrisches Mittel) Beweisen Sie die folgende Ungleichung für reelle Zahlen $a_1, \ldots, a_n \geq 0$:*

$$\sqrt[n]{\prod_{i=1}^n a_i} \leq \frac{1}{n}\sum_{i=1}^n a_i.$$

(a) *Führen Sie den allgemeinen Fall auf den Fall zurück, in dem $\prod_{i=1}^n a_i = 1$ ist.*
(b) *Zeigen Sie mit vollständiger Induktion, dass für reelle Zahlen $a_1, \ldots, a_n \geq 0$ mit $\prod_{i=1}^n a_i = 1$ gilt*

$$\sum_{i=1}^n a_i \geq n.$$

Hinweis: *Fassen Sie im Induktionsschritt für die Anwendung der Induktionsvoraussetzung zwei geeignete Summanden zu einem Produkt zusammen und benutzen Sie die Ungleichung $(a_i - 1)(a_j - 1) \leq 0$.*

Übungsaufgabe 1.6.30. *Archimedes Intervallschachtelung für π*

1.7. Komplexe Zahlen

Definition 1.7.1 (Der Körper der komplexen Zahlen $\mathbb{C}$). *Auf der Menge $\mathbb{R}^2 = \{(a,b) \mid a, b \in \mathbb{R}\}$ definieren wir eine Körperstruktur durch die folgenden Axiome:*

Addition: $(a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2)$
Multiplikation: $(a_1, b_1) \cdot (a_2, b_2) := (a_1a_2 - b_1b_2, a_1b_2 + a_2b_1)$

Dieser Körper wird mit $\mathbb{C}$ bezeichnet und heißt Körper der komplexen Zahlen.

Übungsaufgabe 1.7.2. *Weisen Sie nach, dass für die so definierte Addition und Multiplikation die Körperaxiome erfüllt sind. Zeigen Sie, dass dieser Körper nicht angeordnet werden kann.*

Bemerkung 1.7.3. *Die typische Notation für komplexe Zahlen ist* $a + \mathrm{i}b = (a, b)$*. Für eine komplexe Zahl* $z = a + \mathrm{i}b$ *nennen wir* $\operatorname{Re} z = a$ *den* Realteil *und* $\operatorname{Im} z = b$ *den* Imaginärteil*. Das Element* $\mathrm{i} = (0, 1)$ *heißt* imaginäre Einheit*.*

Bemerkung 1.7.4.

- *Die Idee hier ist, dass wir zu den reellen Zahlen eine neue Zahl hinzufügen wollen, für die* $\mathrm{i}^2 = -1$ *gilt. Eine solche Zahl kann es in* $\mathbb{R}$ *nicht geben (wegen der Anordnung der reellen Zahlen).*
- *Die Regel für die Multiplikation wird in der folgenden Rechnung klarer, in der* $\mathrm{i}^2 = -1$ *benutzt wird:*

$$(a_1 + \mathrm{i}b_1) \cdot (a_2 + \mathrm{i}b_2) = a_1a_2 + \mathrm{i}a_1b_2 + \mathrm{i}a_2b_1 + \mathrm{i}^2b_1b_2 = (a_1a_2 - b_1b_2) + \mathrm{i}(a_1b_2 + a_2b_1).$$

- *Die reellen Zahlen können durch die Abbildung* $\mathbb{R} \to \mathbb{C}\colon a \mapsto (a, 0)$ *als Teilmenge in* $\mathbb{C}$ *realisiert werden. Diese Abbildung ist kompatibel mit Addition und Multiplikation, wir sprechen von einem* Teilkörper*. Dies ist auch die Begründung für die Bezeichnung* Realteil*.*

So wie die reellen Zahlen durch die Zahlengerade veranschaulicht werden können, benutzen wir die *komplexe Zahlenebene* als Veranschaulichung für $\mathbb{C}$. Dafür kann sowohl die Darstellung als $z = x + \mathrm{i}y$ (euklidische Koordinaten) als auch die Darstellung als $z = \|z\| \cdot (\cos \arg(z) + \mathrm{i} \sin \arg(z))$ (Polarkoordinaten) benutzt werden, s. Abbildung 3.

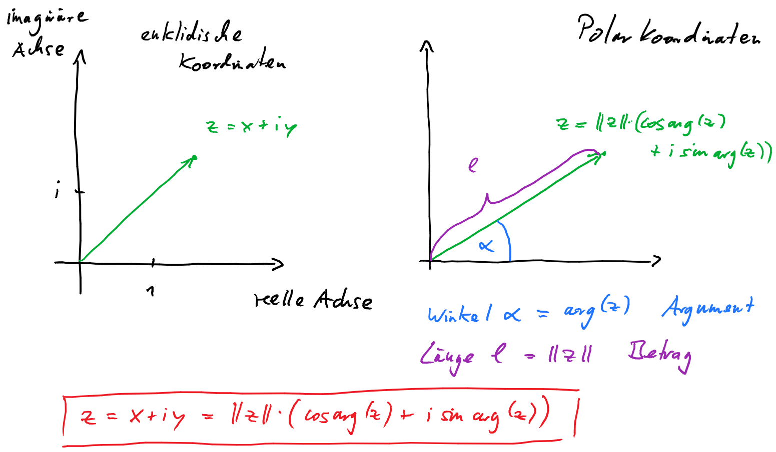


Abbildung 3. Darstellung komplexer Zahlen in der komplexen Zahlenebene

Addition und Multiplikation komplexer Zahlen können ebenfalls gut veranschaulicht werden. Für die Addition gibt es die Parallelogramm-Regel, s. Abbildung 4. Bei der Multiplikation komplexer Zahlen multiplizieren sich die Beträge und die Argumente werden addiert, s. Abbildung 5. Es gibt verschiedene GeoGebra-Veranschaulichungen der Rechenoperationen für komplexe Zahlen, z.B. die Liste unter `https://www.geogebra.org/search/komplexe%20zahlen`.

Beispiel 1.7.5. *In der Transformation zwischen euklidischen und Polarkoordinaten bzw. der geometrischen Beschreibung der Multiplikation komplexer Zahlen sind die Additionstheoreme für Sinus und Kosinus versteckt. Für zwei komplexe Zahlen* $z_1 = \cos \alpha_1 + \mathrm{i} \sin \alpha_1$ *und* $z_2 = \cos \alpha_2 + \mathrm{i} \sin \alpha_2$ *(beide mit Betrag 1) können wir das Produkt auf zwei verschiedene Arten berechnen:*

$$\begin{aligned} z_1 \cdot z_2 &= (\cos \alpha_1 + \mathrm{i} \sin \alpha_1)(\cos \alpha_2 + \mathrm{i} \sin \alpha_2) \\ &= (\cos \alpha_1 \cos \alpha_2 - \sin \alpha_1 \sin \alpha_2) + \mathrm{i}(\cos \alpha_1 \sin \alpha_2 + \cos \alpha_2 \sin \alpha_1) \\ z_1 \cdot z_2 &= \cos(\alpha_1 + \alpha_2) + \mathrm{i} \sin(\alpha_1 + \alpha_2). \end{aligned}$$

□

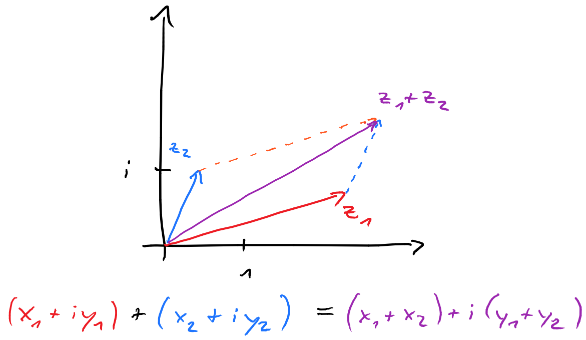


Abbildung 4. Parallelogramm-Regel für Addition komplexer Zahlen

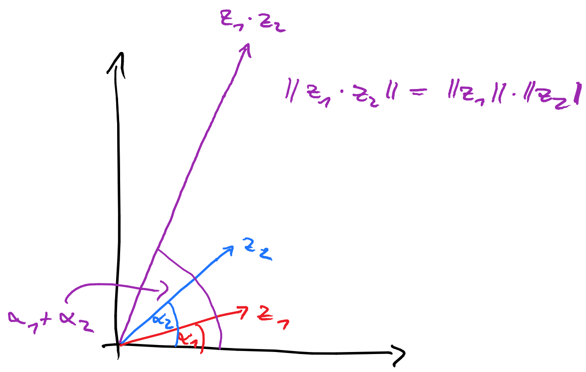


Abbildung 5. Multiplikation komplexer Zahlen in Polarkoordinaten

Definition 1.7.6. *Die* komplexe Konjugation *ist die Abbildung* $\mathbb{C} \to \mathbb{C}\colon z = a + \mathrm{i}b \mapsto \overline{z} = a - \mathrm{i}b$.

Bemerkung 1.7.7.
- *In der komplexen Ebene entspricht komplexe Konjugation der Spiegelung an reeller Achse.*
- *Komplexe Konjugation ist kompatibel mit Addition und Multiplikation, d.h. es gelten*

$$\begin{aligned} \overline{z_1 + z_2} &= \overline{z_1} + \overline{z_2} \\ \overline{z_1 \cdot z_2} &= \overline{z_1} \cdot \overline{z_2} \end{aligned}$$

Definition 1.7.8. *Auf den komplexen Zahlen haben wir einen Absolutbetrag, der wie folgt definiert ist:*

$$\|\cdot\|\colon \mathbb{C} \to \mathbb{R}_{\geq 0}\colon z = a + \mathrm{i}b \mapsto \|z\| = \sqrt{z \cdot \overline{z}} = \sqrt{a^2 + b^2}.$$

Bemerkung 1.7.9.
- *In der komplexen Zahlenebene beschreibt der Absolutbetrag einer komplexen Zahl deren Abstand zum Ursprung.*
- *Es gelten dieselben Rechenregeln wie für den Absolutbetrag der reellen Zahlen. (Achtung: Die komplexen Zahlen sind kein angeordneter Körper!)*
- *Mit Konjugation und Absolutbetrag haben wir für* $z \in \mathbb{C} \setminus \{0\}$

$$\frac{1}{z} = \frac{\overline{z}}{\|z\|}$$

Geometrisch ist das Invertieren komplexer Zahlen zusammengesetzt aus einer Spiegelung am Einheitskreis $\mathrm{S}^1 = \{z \in \mathbb{C} \mid \|z\| = 1\}$ *und einer Spiegelung an der reellen Achse.*

Proposition 1.7.10. *Der Absolutbetrag hat die folgenden Eigenschaften:*

positive Definitheit: $\|x\| \geq 0$ *für alle* $x \in \mathbb{C}$,
und $\|x\| = 0$ *genau dann, wenn* $x = 0$.
Multiplikativität: $\|x \cdot y\| = \|x\| \cdot \|y\|$ *für alle* $x, y \in \mathbb{C}$.
Dreiecks-Ungleichung: $\|x + y\| \leq \|x\| + \|y\|$ *für alle* $x, y \in \mathbb{C}$.

Beweis. Übungsaufgabe □

Proposition 1.7.11. *In* $\mathbb{C}$ *hat jede quadratische Gleichung* $ax^2 + bx + c = 0$ *eine Lösung.*

Beweis. Wir nehmen an, dass $a \neq 0$. (Andernfalls haben wir eine viel einfachere lineare Gleichung.)

Wir beweisen die Aussage mit der Lösungsformel für quadratische Gleichungen:

$$x_{1/2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Äquivalent dazu ist die Aussage, dass die Lösungen von $x^2 + px + q = 0$ durch

$$x_{1/2} = \frac{-p}{2} \pm \sqrt{\frac{p^2}{4} - q}$$

gegeben sind. Die Äquivalenz sehen wir nach Division von $ax^2 + bx + c = 0$ durch a, mit $p = \frac{b}{a}$ und $q = \frac{c}{a}$.

Zur Herleitung der Lösungsformel benutzen wir die quadratische Ergänzung: nach der binomischen Formel ist

$$(x + r)^2 = x^2 + 2rx + r^2.$$

Wir vergleichen Koeffizienten mit $x^2 + px + q$ und setzen $r = \frac{p}{2}$. Dann ist

$$x^2 + px + q = x^2 + px + q + \left(\frac{p}{2}\right)^2 - \left(\frac{p}{2}\right)^2 = \left(x + \frac{p}{2}\right)^2 + q - \left(\frac{p}{2}\right)^2 = 0.$$

Daraus folgt

$$\left(x + \frac{p}{2}\right)^2 = \frac{p^2}{4} - q,$$

Wir müssen also nur noch die Gleichung $y^2 = d$ für $d \in \mathbb{C}$ lösen. Wenn $d \in \mathbb{R}$ ist, haben wir für $d > 0$ die Existenz von Wurzeln in $\mathbb{R}$ nach Satz 1.6.25. Tatsächlich haben wir genau zwei Lösungen $y = \pm\sqrt{d}$ in $\mathbb{R}$ (wegen Eindeutigkeit der positiven Wurzeln). Für $d = 0$ haben wir die Lösung $y = 0$. Für $d < 0$ setzen wir $\sqrt{d} = \mathrm{i}\sqrt{-d}$ und benutzen wieder den reellen Fall. Auch hier gibt es dann wieder zwei Lösungen $y = \pm\sqrt{d} = \pm\mathrm{i}\sqrt{-d}$. Wenn $d = \|d\|(\cos\alpha + \mathrm{i}\sin\alpha)$ komplex ist, haben wir als Lösungen

$$y = \pm\sqrt{\|d\|}\left(\cos\frac{\alpha}{2} + \mathrm{i}\sin\frac{\alpha}{2}\right).$$

Damit folgern wir aus $\left(x + \frac{p}{2}\right)^2 = \frac{p^2}{4} - q$ die Lösungsformel:

$$x_{1/2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}$$

□

Definition 1.7.12. *n-te Einheitswurzeln Für* $n \in \mathbb{N}_{>0}$ *definieren wir die* n-te Einheitswurzel $\zeta_n := \cos\frac{2\pi}{n} + \mathrm{i}\sin\frac{2\pi}{n}$.

BEMERKUNG 1.7.13. *Es gilt* $\zeta_n^n = 1$. *Die Potenzen* ζ_n^i *bilden die Ecken eines regelmäßigen* n*-Ecks, dessen Ecken auf dem Einheitskreis liegen.* ζ_3, $\zeta = \mathrm{i}$.

ÜBUNGSAUFGABE 1.7.14. *(a) Schreiben Sie die folgenden komplexen Zahlen in der Form* $x + \mathrm{i}y$*:*

$$\mathrm{i}^4 + \mathrm{i}^5 + \mathrm{i}^6 + \mathrm{i}^7, \qquad \frac{\mathrm{i}}{1-\mathrm{i}}, \qquad \frac{1}{\mathrm{i}} + \frac{3}{1+\mathrm{i}}, \qquad (1+\mathrm{i})^n + (1-\mathrm{i})^n \text{ für } n \in \mathbb{N}.$$

(b) Ermitteln Sie die komplexe Zahl z*, die die Gleichung* $\frac{2+3\mathrm{i}}{2} z + \frac{5+2\mathrm{i}}{1+\mathrm{i}} = 8 + 2\mathrm{i}$ *erfüllt.*
(c) Ermitteln Sie die komplexe Quadratwurzel aus $2(-1 + \mathrm{i}\sqrt{3})$.
(d) Zeichnen Sie die folgenden Mengen in der komplexen Zahlenebene

$$\begin{aligned} M_1 &= \{z \in \mathbb{C} \mid 1 < \|z - \mathrm{i}\| \leq 2\} \\ M_2 &= \{z \in \mathbb{C} \mid 1 < \mathrm{i}(z - \overline{z}) < 2\} \\ M_3 &= \{z \in \mathbb{C} \mid \|z + 1\| \leq \|z - 2\|\} \end{aligned}$$

Hinweis: *Bei einigen Teilaufgaben hilft die geometrische Veranschaulichung beim Finden der Lösung.*

ÜBUNGSAUFGABE 1.7.15. *Gegeben seien die komplexen Zahlen* $z_1 = 1 + \mathrm{i}$ *und* $z_2 = 2 - 3\mathrm{i}$. *Schreiben Sie die folgenden komplexen Zahlen in der Form* $a + b\mathrm{i}$*:*

$$\overline{z_j},\ -z_j,\ z_j\overline{z_j},\ \frac{1}{z_j},\ z_j - \overline{z_j} \text{ jeweils für } j = 1,2; \qquad \frac{z_1}{z_1 + z_2}, \frac{z_1}{\overline{z_2} - z_1^2}.$$

ÜBUNGSAUFGABE 1.7.16. *Zeichnen Sie die folgenden Mengen von komplexen Zahlen in der komplexen Zahlenebene*

(1) $\|z\| \geq \frac{15}{\|z\| - 2}$
(2) $\|z\| = |\mathrm{Re}(z)| + |\mathrm{Im}(z)|$
(3) $\|z\|^2 \leq 2\mathrm{Re}(z)$
(4) $|\mathrm{Re}(z)| + |\mathrm{Im}(z)| \leq 4$

ÜBUNGSAUFGABE 1.7.17. *"Auf der Schatzinsel ist ein kleiner Baum* B_1 *und ein großer Baum* B_2*, sowie ein Kreuz* X*. Gehe vom Kreuz nach* B_1 *und noch einmal genauso lang weiter und um dieselbe Strecke nach links (*$90°$*) um die Ecke. Markiere diese Stelle* M_1*. Geht man vom Baum* B_2 *zum Kreuz* X *und links um* $90°$ *dieselbe Strecke verlängert, ergibt dies einen zweiten Punkt* M_2*. Der Schatz ist in der Mitte zwischen* M_1 *und* M_2 *vergraben."*

Behauptung: *Das Kreuz* X *ist verschwunden, aber der Schatz lässt sich trotzdem finden.*

Hinweis: *Benutzen Sie die komplexen Zahlen.*

ÜBUNGSAUFGABE 1.7.18. *Seien* $a = m^2 + n^2$ *und* $b = p^2 + q^2$ *Summen von je zwei Quadraten mit* $m, n, p, q \in \mathbb{Z}$*. Zeigen Sie, dass auch* ab *eine Summe von zwei Quadraten ganzer Zahlen ist.*

1.8. Reell- und komplexwertige Funktionen

Wir erinnern uns an den Abschnitt zu Abbildungen von Mengen: eine Abbildung $f\colon A \to B$ ordnet jedem Element aus A genau ein Element aus B zu:

$$f\colon A \to B\colon a \mapsto f(a)$$

Dabei heißt A *Definitionsbereich* und $f(A) \subseteq B$ Wertebereich. Für die Vorlesung Analysis 1 betrachten wir die folgenden Abbildungen, die wir auch *Funktionen* nennen:

DEFINITION 1.8.1. *Eine* reell-wertige *(bzw.* komplex-wertige*) Funktion auf einer Menge X ist eine Abbildung $f\colon X \to \mathbb{R}$ (bzw. $f\colon X \to \mathbb{C}$). Ist $X \subseteq \mathbb{R}$ sprechen wir von einer Funktion in einer reellen Variablen (oder Veränderlichen), ist $X \subseteq \mathbb{C}$ sprechen wir von einer Funktion in einer komplexen Veränderlichen.*

BEMERKUNG 1.8.2.
- *Zur Veranschaulichung einer Funktion $f\colon U \to \mathbb{R}$ mit $U \subseteq \mathbb{R}$ benutzen wir typischerweise Zeichnungen des Graphen $\Gamma(f) = \{(x, f(x)) \mid x \in \mathbb{R}\} \subseteq \mathbb{R}^2$ in der Ebene.*
- *In der Analysis 2 wird es dann um Funktionen $f\colon U \to \mathbb{R}^m$ mit $U \subseteq \mathbb{R}^n$ gehen.*

BEISPIEL 1.8.3.
- *Für $a \in \mathbb{C}$ haben wir die konstante Funktion $X \to \mathbb{C}\colon x \mapsto a$.*
- *Ein Beispiel für eine Funktion (deren Graph nicht wirklich gezeichnet werden kann) ist*
$$\mathbb{R} \to \mathbb{R}\colon x \mapsto \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$
Solche Abbildungen kommen in der Analysis 1 nur sehr selten vor
- *Wir betrachten Körper im freien Fall. In dieser Situation bekommen wir zwei Funktionen: eine Funktion erhalten wir durch Interpolation aus einer Messreihe (im Experiment messen wir den zurückgelegten Weg in Abhängigkeit von der Zeit). Eine andere Funktion, die den zurückgelegten Weg in Abhängigkeit von der Zeit beschreibt, erhalten wir abgeleitet aus den Gravitationsgesetzen: $s(t) = \frac{gt^2}{2}$, wobei $g \simeq 9.81\frac{m}{s^2}$ die Fallbeschleunigung ist. Die Qualität der Naturgesetze bemisst sich hier daran, wie gut die gemessene und die aus den Naturgesetzen abgeleitete Funktion übereinstimmen.*
- *Die Temperaturverteilung in einem Raum (in Abhängigkeit von der Zeit) bei Wärmeeinstrahlung können wir als Funktion $T\colon U \to \mathbb{R}\colon (x, y, z, t) \mapsto T(x, y, z, t)$ mit $U \subseteq \mathbb{R}^4$ (drei Raumkoordinaten x, y, z und eine Zeitkoordinate t) auffassen. Solche Situationen werden in der Analysis 2 betrachtet.*
- *Die Gaußklammer, s. Abbildung:*
$$\lfloor - \rfloor\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto \lfloor x \rfloor := \max\{k \in \mathbb{Z} \mid k \leq x\}$$
- *Allgemeiner haben wir sogenannte* Treppenfunktionen *(oder* stückweise konstante Funktionen*). Für reelle Zahlen $a < b$ unterteilen wir das Intervall $[a, b]$ in kleinere Intervalle $a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$ und definieren eine Funktion*
$$[a, b] \to \mathbb{R}\colon x \mapsto \begin{cases} c_i & x \in (t_i, t_{i+1}) \\ d_i & x = t_i, \end{cases}$$
wobei $c_0, \ldots, c_{n-1}, d_0, \ldots, d_n \in \mathbb{R}$ reelle Zahlen sind, s. Abbildung.
- *Für rationale Zahlen $\frac{p}{q}$ mit $q \neq 0$ können wir die Potenz einer reellen Zahl $x \in \mathbb{R}$ mit rationalen Exponenten definieren:*
$$x^{\frac{p}{q}} := \sqrt[q]{x^p}$$
Dies liefert die Funktionen
$$(-)^{\frac{p}{q}}\colon \mathbb{R}_{\geq 0} \to \mathbb{R}\colon x \mapsto \sqrt[q]{x^p}.$$
- *Im zweiten Teil des Abschnitts werden insbesondere Polynomfunktionen diskutieren:*
$$f\colon \mathbb{C} \to \mathbb{C}\colon x \mapsto \sum_{i=0}^{n} a_i x^i, \quad a_i \in \mathbb{C}.$$

□

DEFINITION 1.8.4 (Operationen mit Funktionen). *Seien $f, g\colon X \to \mathbb{C}$ zwei Funktionen. Wir definieren*

Summe: $f + g\colon X \to \mathbb{C}\colon x \mapsto (f+g)(x) := f(x) + g(x)$

Produkt: $f \cdot g\colon X \to \mathbb{C}\colon x \mapsto (f \cdot g)(x) := f(x) \cdot g(x)$

Quotient: *Wenn $g \neq 0$ ist, definieren wir*

$$\frac{f}{g}\colon \{x \in X \mid g(x) \neq 0\} \to \mathbb{C}\colon x \mapsto \frac{f}{g}(x) := \frac{f(x)}{g(x)}$$

Speziell für komplexwertige Funktionen können wir auch noch definieren:

Konjugation: $\overline{f}\colon X \to \mathbb{C}\colon x \mapsto \overline{f}(x) := \overline{f(x)}$

Real- und Imaginärteil: $\mathrm{Re}(f)\colon X \to \mathbb{C}\colon x \mapsto (\mathrm{Re}(f))(x) := \mathrm{Re}(f(x))$, *analog für Imaginärteil.*

Speziell für reell-wertige Funktionen $f, g\colon X \to \mathbb{R}$ können wir definieren:

Betrag: $|f|\colon X \to \mathbb{R}\colon x \mapsto |f|(x) := |f(x)|$

Maximum, Minimum:

$$\max(f, g)\colon X \to \mathbb{R}\colon x \mapsto (\max(f,g))(x) := \max(f(x), g(x)),$$

analog für Minimum.

BEMERKUNG 1.8.5. *Alle obigen Definition erhalten wir daraus, dass wir Operationen, die wir für reelle oder komplexe Zahlen haben (Addition, Multiplikation, Konjugation, Minimum, etc) auf die Werte der Funktionen f, g anwenden. Mit Addition und Multiplikation bilden die Funktionen einen kommutativen Ring, d.h. alle Körperaxiome bis auf die Existenz von multiplikativen Inversen sind erfüllt.*

DEFINITION 1.8.6 (Komposition).

- *Zwei komplex-wertige Funktionen $f\colon X \to \mathbb{C}$ und $g\colon Y \to \mathbb{C}$ können zusammengesetzt werden, wenn $f(X) \subseteq Y \subseteq \mathbb{C}$ ist. Dann haben wir (wie in der Definition der Komposition von Abbildungen)*

$$(g \circ f)(x) := g(f(x))$$

- *Ein Spezialfall ist die sogenannte* Einschränkung *von Funktionen. Für eine Funktion $f\colon X \to \mathbb{C}$ und eine Teilmenge $A \subseteq X$ definieren wir die Einschränkung von X auf A durch*

$$f|_A\colon A \to \mathbb{C}\colon a \in A \mapsto f(a).$$

Dies ist die Komposition der Inklusionsabbildung $i\colon A \to X\colon a \in A \mapsto a \in X$ und der Abbildung f.

BEMERKUNG 1.8.7. *Für $f\colon X \to \mathbb{C}$ eine injektive Funktion mit $X \subseteq \mathbb{C}$ existiert eine* Umkehrfunktion *$f^{-1}\colon f(X) \to X \subseteq \mathbb{C}$. Eine analoge Aussage gilt für reellwertige Funktionen. Geometrisch erhalten wir den Graphen einer Umkehrfunktion für $f\colon X \subseteq \mathbb{R} \to \mathbb{R}$ durch Spiegelung an der Diagonalen $(x, x) \in \mathbb{R}^2$ in der Ebene.*

1.9. Polynome und rationale Funktionen

DEFINITION 1.9.1.

- Polynomfunktionen *sind Funktionen der Form*

$$f\colon K \to K\colon x \mapsto f(x) = \sum_{i=0}^{n} a_i x^i,$$

mit $K = \mathbb{R}$ oder $K = \mathbb{C}$. Dabei sind die $a_i \in K$ die Koeffizienten.

- *Der* Grad *eines Polynoms f ist definiert als $\deg f = \max\{i \mid a_i \neq 0\}$, der höchste von Null verschiedene Koeffizient $a_{\deg f}$ heißt* Leitkoeffizient.

Bemerkung 1.9.2.
- *Das Polynom $f = 0$ mit $a_i = 0$ für alle i heißt* Nullpolynom.
- *Für Summe und Produkt von Polynomen haben wir (mit Induktion)*

$$\begin{aligned}\left(\sum_{i=0}^{n} a_i x^i\right) + \left(\sum_{i=0}^{m} b_i x^i\right) &= \sum_{i=0}^{\max(n,m)} (a_i + b_i) x^i \\ \left(\sum_{i=0}^{n} a_i x^i\right) \cdot \left(\sum_{i=0}^{m} b_i x^i\right) &= \sum_{i=0}^{n+m} \sum_{p+q=i} a_p b_q x^i\end{aligned}$$

Die Menge der Polynome mit reellen bzw. komplexen Koeffizienten in der Variablen x wird mit $\mathbb{R}[x]$ bzw. $\mathbb{C}[x]$ bezeichnet. Mit Addition und Multiplikation haben wir wieder kommutative Ringe, d.h. alle Körperaxiome bis auf die Existenz von multiplikativen Inversen sind erfüllt.
- *Für den Grad von Polynomen gilt*

$$\begin{aligned}\deg(f \cdot g) &= \deg f + \deg g \\ \deg(f + g) &\leq \max(\deg f, \deg g)\end{aligned}$$

Übungsaufgabe 1.9.3. *Zeigen Sie, dass ein Polynom $0 \neq f \in \mathbb{C}[x]$ genau dann ein multiplikatives Inverses hat, wenn es konstant ist.*

Proposition 1.9.4 (Division mit Rest). *Sei $g \neq 0$ ein Polynom. Dann existieren zu jedem Polynom f eindeutige Polynome q und r so dass*

(1) $\deg r < \deg g$ *oder* $r = 0$
(2) $f = q \cdot g + r$

Beweis. Sei $g = \sum_{i=0}^{m} b_i x^i$ das gegebene Polynom mit Leitkoeffizient $b_m \neq 0$. Wir beweisen die Existenzaussage mit Induktion über den Grad des Polynoms. Der Induktionsanfang ist der Fall $\deg f < \deg g$, in dem $q = 0$ und $r = f$ die Bedingungen erfüllen.

Für den Induktionsschritt nehmen wir an, dass die Aussage für alle Polynome f mit $\deg f < n$ existiert. Sei $f = \sum_{i=0}^{n} a_i x^i$ ein Polynom vom Grad n. Dann ist

$$f_1 := f - a_n b_m^{-1} x^{n-m} g$$

ein Polynom mit $\deg f_1 < n$, also existiert nach Induktionsvoraussetzung eine Zerlegung $f_1 = q_1 \cdot g + r$ mit $r = 0$ oder $\deg r < \deg g$. Dann ist

$$f = q_1 \cdot g + r + a_n b_m^{-1} x^{n-m} g = (q_1 + a_n b_m x^{n-m}) g + r,$$

und wir erhalten eine Zerlegung von f mit $q = q_1 + a_n b_m x^{n-m}$ und r. Nach dem Prinzip der vollständigen Induktion haben wir die Existenz der Polynome q und r mit den gewünschten Eigenschaften bewiesen.

Für die Eindeutigkeitsaussage sei $f = q' \cdot g + r'$ eine weitere Zerlegung mit $q \neq q'$. Dann ist $(q' - q)g = r - r'$, aber dann ist nach Voraussetzung $\deg(q' - q)g = \deg(r - r') < \deg g$. Auf der anderen Seite gilt für $q' - q \neq 0$ immer $\deg(q' - q)g = \deg(q' - q) + \deg g \geq \deg g$, und wir erhalten einen Widerspruch. Also sind die Polynome q und r eindeutig. □

Definition 1.9.5. *Für ein Polynom $f(x) = \sum_{i=0}^{n} a_i x^i$ nennen wir ein α mit $f(\alpha) = 0$ eine* Nullstelle *von f.*

Korollar 1.9.6 (Abspaltung von Linearfaktoren). *Ist α eine Nullstelle von f, so ist f durch $x - \alpha$ teilbar, d.h. $f(x) = (x - \alpha) q(x)$ für ein Polynom q mit* $\deg q = \deg f - 1$.

Beweis. Eine Polynomdivision mit $g = (x - \alpha)$ liefert $f(x) = q(x)(x - \alpha) + r(x)$ mit $\deg r < \deg g = 1$. Also ist $r = 0$ oder $\deg r = 0$ (also ist r konstant). Auswertung der Gleichung an $x = \alpha$ liefert

$$0 = f(\alpha) = \underbrace{q(\alpha)(\alpha - \alpha)}_{0} + r(\alpha).$$

Daraus folgt $r = 0$ wie behauptet. □

Definition 1.9.7. *Für ein Polynom $f(x) = \sum_{i=0}^n a_i x^i \neq 0$ nennen wir $\alpha \in \mathbb{C}$ eine k-*fache Nullstelle* von f, wenn f durch $(x - \alpha)^k$ aber nicht durch $(x - \alpha)^{k+1}$ teilbar ist. Wir nennen k die* Vielfachheit *der Nullstelle α.*

Korollar 1.9.8. *Ein Polynom $\neq 0$ vom Grad n hat höchstens n Nullstellen.*

Beweis. Mit Induktion über den Grad: der Induktionsanfang ist, das ein konstantes Polynom $\neq 0$ keine Nullstelle hat. Für den Induktionsschritt benutzen wir die Abspaltung von Linearfaktoren aus Korollar 1.9.6. □

Satz 1.9.9 (Identitätssatz). *Stimmen die Werte der beiden Polynome $f(x) = \sum_{i=0}^n a_i x^i$ und $g(x) = \sum_{i=0}^n b_i x^i$ an $n+1$ verschiedenen Stellen überein, dann gilt schon $a_i = b_i$ für alle $i \in \{0, \dots, n\}$ und damit $f(x) = g(x)$ für alle $x \in \mathbb{C}$.*

Beweis. Die Differenz $f - g$ hat $n+1$ verschiedene Nullstellen und $\deg(f-g) \leq n$. Damit folgt aus Korollar 1.9.8, dass $f - g = 0$ sein muss. □

Satz 1.9.10 (Fundamentalsatz der Algebra). *Jedes nicht konstante Polynom mit komplexen Koeffizienten hat mindestens eine komplexe Nullstelle.*

Bemerkung 1.9.11. *Für diesen wichtigen Satz gibt es viele Beweise (mit Methoden aus Analysis, Funktionentheorie, Topologie, ...). Genauer können wir sagen, dass ein komplexes Polynom vom Grad n genau n Nullstellen hat, wenn wir mit Vielfachheiten wie in Definition 1.9.7 zählen.*

Korollar 1.9.12 (Linearfaktorzerlegung). *Jedes Polynom $f \in \mathbb{C}[z]$ vom Grad n kann als*

$$f(z) = a(z - \alpha_1)^{m_1} \cdots (z - \alpha_s)^{m_s}$$

mit $\sum m_i = n$ geschrieben werden.

Beweis. Der Beweis ist eine Kombination aus dem Fundamentalsatz der Algebra 1.9.10 und der Abspaltung von Linearfaktoren aus Korollar 1.9.6. □

Übungsaufgabe 1.9.13. *(a) Zeigen Sie, dass für ein Polynom $P(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ der Koeffizient a_0 bis auf Vorzeichen gleich dem Produkt der komplexen Nullstellen des Polynoms ist.*

(b) Zerlegen Sie das Polynom $P(z) = z^4 - 2z^3 + 5x^2 - 8x + 4$ in Linearfaktoren.

Korollar 1.9.14 (Zerlegung reeller Polynome). *Jedes reelle Polynom $f(x) = \sum_{i=0}^n a_i x^i$ vom Grad n hat eine Faktorisierung der Form*

$$f(x) = a_n(x - b_1)^{l_1} \cdots (x - b_r)^{l_r}(x^2 + c_1 x + d_1)^{k_1} \cdots (x^2 + c_s x + d_s)^{k_s}$$

mit reellen Nullstellen $b_i \in \mathbb{R}$ mit Vielfachheit k_i und quadratischen Polynomen $x^2 + c_i x + d_i$, die keine reellen Nullstellen haben.

Beweis. Wir fassen f als komplexes Polynom auf. Für eine komplexe Nullstelle $w \in \mathbb{C}$ ist dann $\overline{w}$ auch eine Nullstelle:

$$f(\overline{w}) = \sum_{i=0}^n a_i \overline{w}^i = \overline{\sum_{i=0}^n a_i w^i} = \overline{f(w)} = 0.$$

Das zweite Gleichheitszeichen benutzt dabei, dass $a_i \in \mathbb{R}$ für alle i ist. Damit sind also die komplexen Nullstellen von f entweder reell oder treten in komplex-konjugierten Paaren auf.

Wir betrachten die Linearfaktorzerlegung aus Korollar 1.9.12:

$$f(z) = a_n(z - w_1)^{l_1} \cdots (z - w_k)^{l_k}.$$

Nach den vorigen Überlegungen ist entweder $w_i \in \mathbb{R}$ oder wir haben

$$(z - w_i)^{l_i}(z - \overline{w}_i)^{l_i} = (z^2 - (w + \overline{w})z + w\overline{w})^{l_i} = (z^2 - 2\operatorname{Re}(w) \cdot z + \|w\|^2)^{l_i}.$$

Dies liefert die behauptete Zerlegung. □

Übungsaufgabe 1.9.15. *Gegeben seien $n+1$ verschiedene Stellen $z_0, \ldots, z_n \in \mathbb{C}$ und $n+1$ beliebige Werte $w_0, \ldots, w_n \in \mathbb{C}$. Zeigen Sie, dass es eindeutige Zahlen $c_0, \ldots, c_n \in \mathbb{C}$ gibt, so dass für das Polynom*

$$P(z) = c_0 + \sum_{k=0}^{n-1} c_{k+1}(z - z_0) \cdots (z - z_k)$$

gilt $P(z_i) = w_i$ für alle $i \in \{0, \ldots, n\}$.

Definition 1.9.16. Rationale Funktionen *sind Quotienten von Polynomen:*

$$f(x) = \frac{p(x)}{q(x)} = \frac{\sum_{i=0}^n a_i x^i}{\sum_{j=0}^m b_j x^j}, \quad a_n, b_m \neq 0$$

Bemerkung 1.9.17.
- *Dabei können Linearfaktoren für gemeinsame Nullstellen gekürzt werden.*
- *Wenn p, q keine gemeinsamen Nullstellen haben, dann ist der maximale Definitionsbereich der rationalen Funktion $D(f) = \{x \in \mathbb{C} \mid q(x) \neq 0\}$.*
- *Eine k-fache Nullstelle von $q(x)$ führt zu einem* Pol k-ter Ordnung *für f.*

Übungsaufgabe 1.9.18. *Zeigen Sie, dass rationale Funktionen mit reellen Koeffizienten einen Körper bilden. Dieser Körper wird mit $\mathbb{R}(x)$ bezeichnet.*

KAPITEL 2

Konvergenz für Folgen und Reihen

2.1. Grenzwerte: Definition und Beispiele

DEFINITION 2.1.1. *Eine* Folge *reeller Zahlen ist eine Abbildung* $a\colon \mathbb{N} \to \mathbb{R}$. *Analog ist eine Folge komplexer Zahlen eine Abbildung* $a\colon \mathbb{N} \to \mathbb{C}$.

Notation: *Wir schreiben meist* a_n *für den Wert* $a(n)$, *und* $(a_n)_{n\in\mathbb{N}}$ *für die Folge.*

BEISPIEL 2.1.2.
(1) *Das einfachste Beispiel für eine Folge ist die* konstante Folge $a_n = c$.
(2) *Ein Beispiel für eine* alternierende Folge *ist* $a_n = (-1)^n$.
(3) *arithmetische Folge:* $a_n = a_0 + nd$
(4) *Für* $a_0, q \in \mathbb{R}$ *haben wir die* geometrische Folge $a_n = a_0 q^n$.
(5) *Die* Fibonacci-Folge *ist ein Beispiel für eine* rekursiv definierte Folge*:* $a_0 = a_1 = 1$ *und* $a_n = a_{n-1} + a_{n-2}$ *für* $n \geq 2$.

□

DEFINITION 2.1.3. *Eine Folge* (a_n) *reeller Zahlen heißt* konvergent, *wenn es eine Zahl* $a \in \mathbb{R}$ *gibt, so dass für jede reelle Zahl* $\epsilon > 0$ *eine natürliche Zahl* $N = N(\epsilon)$ *existiert, so dass für alle natürlichen Zahlen* $n > N$ *gilt* $|a_n - a| < \epsilon$. *Die Zahl* a *(wenn sie existiert) heißt* Grenzwert *(oder* Limes*) der Folge.*

Notation: $\lim_{n\to\infty} a_n = a$ *oder* $a_n \to a$ *für* $n \to \infty$.

Analog definieren wir die Konvergenz einer Folge (a_n) komplexer Zahlen mit der Bedingung $\|a_n - a\| < \epsilon$.

BEMERKUNG 2.1.4.
- *In mathematischer Notation*
$$\forall \epsilon > 0\ \exists N \in \mathbb{N}\ \forall n > N\colon |a_n - a| < \epsilon.$$
Die Negation dieser Aussage ist
$$\exists \epsilon > 0\ \forall N \in \mathbb{N}\ \exists n > N\colon |a_n - a| \geq \epsilon.$$
Das bedeutet, dass eine Folge divergiert *(d.h.* nicht konvergiert*), wenn es einen Abstand* ϵ *gibt, so dass Folgeglieder* a_n *mit beliebig großem Index* n *existieren, die von* a *mindestens Abstand* ϵ *haben.*
- *Wir sagen, dass eine Folge* $(a_n)_{n\in\mathbb{N}}$ bestimmt gegen $+\infty$ (bzw. $-\infty$) divergiert, *wenn für alle reellen Zahlen* $b \in \mathbb{R}$ *ein* $N \in \mathbb{N}$ *existiert, so dass für alle* $n > N$ *gilt* $a_n > b$ *(bzw.* $a_n < b$*). Dafür schreiben wir auch* $\lim_{n\to\infty} a_n = \infty$ *(bzw.* $\lim_{n\to\infty} a_n = -\infty$*).*
- *Unter einer* (offenen) ϵ-Umgebung $U_\epsilon(x)$ *einer Zahl* x *verstehen wir die Menge aller Zahlen, die von* x *Abstand kleiner als* ϵ *haben. Für die reellen Zahlen ist dies ein offenes Intervall:* $U_\epsilon(x) = \{y \in \mathbb{R} \mid |x - y| < \epsilon\} = (x - \epsilon, x + \epsilon) \subseteq \mathbb{R}$. *Für die komplexen Zahlen ist dies eine Kreisscheibe (ohne Rand!) um* x *in der Zahlenebene* $U_\epsilon(x) = \{y \in \mathbb{C} \mid \|x - y\| < \epsilon\}$. *Geometrisch interpretiert bedeutet Konvergenz (in beiden Fällen), dass für beliebig kleine Abstände* ϵ fast alle *(d.h. alle bis auf endlich viele) Folgenglieder in der* ϵ*-Umgebung des Grenzwerts liegen.*
- *Eine Folge, die gegen* 0 *konvergiert, heißt auch* Nullfolge.

Beispiel 2.1.5. *Wir diskutieren das Konvergenzverhalten der geometrischen Folge $a_n = q^n$.*

- *Wenn $q > 1$ ist, existiert für jedes $b \in \mathbb{R}$ ein $n \in \mathbb{N}$, so dass $q^n > b$ ist, s. Proposition 1.5.12, Fall (2). Die Folge divergiert also (tatsächlich divergiert die Folge bestimmt gegen $+\infty$).*
- *Für $q = 1$ ist die Folge konstant und $\lim_{n\to\infty} q^n = 1$.*
- *Für $0 < q < 1$ haben wir $\lim_{n\to\infty} q^n = 0$. Nach Definition bedeutet das, dass für alle $\epsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass für alle $n > N$ gilt $q^n < \epsilon$. Dies folgt aus Proposition 1.5.12, Fall (3).*
- *Für $q = 0$ haben wir wieder eine konstante Folge und $\lim_{n\to\infty} q^n = 0$.*
- *Für $-1 < q < 0$ haben wir $\lim_{n\to\infty} q^n = 0$, wie im Fall $0 < q < 1$.*
- *Für $q = -1$ ist $a_n = (-1)^n$. Für jedes $N \in \mathbb{N}$ existiert ein $n > N$ so dass $a_n = q^n = 1$ und $a_{n+1} = q^{n+1} = -1$. Damit ist für jede Wahl von a immer $\max(|a_n - a|, |a_{n+1} - a|) \geq \frac{1}{2}$, also konvergiert die Folge nicht.*
- *Für $q < -1$ konvergiert die Folge nicht. Die Beträge $|q^n|$ wachsen unbeschränkt, allerdings gibt es noch einen qualitativen Unterschied zum Fall $q > 1$: das Vorzeichen der Folgeglieder $a_n = q^n$ alterniert, so dass wir keine bestimmte Divergenz haben.*

□

Lemma 2.1.6. *Der Grenzwert einer konvergenten Folge ist eindeutig.*

Beweis. Seien $a \neq a'$ zwei Zahlen, die die Grenzwertbedingung erfüllen. Für $\epsilon = \frac{|a'-a|}{3}$ existieren dann $N, N' \in \mathbb{N}$, so dass

$$\forall n > N\colon |a_n - a| < \epsilon \quad \text{und} \quad \forall n > N'\colon |a_n - a'| < \epsilon$$

gilt. Für $n > \max(N, N')$ folgt

$$|a - a'| = |a - a_n + a_n - a'| \leq |a - a_n| + |a_n - a'| < \frac{2}{3}|a' - a|.$$

□

Liste wichtiger Grenzwerte

Proposition 2.1.7. *(1) $\lim_{n\to\infty} \frac{1}{n^s} = 0$ für alle $s \in \mathbb{Q}$ mit $s > 0$.*
(2) $\lim_{n\to\infty} \sqrt[n]{a} = 1$ für alle $a \in \mathbb{R}$ mit $a > 0$.
(3) $\lim_{n\to\infty} \sqrt[n]{n} = 1$
(4) $\lim_{n\to\infty} q^n = 0$ für $q \in \mathbb{C}$ mit $\|q\| < 1$.

Beweis. (1) Für $\epsilon > 0$ setzen wir $N = \epsilon^{-\frac{1}{s}}$. Dann gilt für $n > N$ auch $n^s > N^s$ und damit $|\frac{1}{n^s}| < \frac{1}{N^s} = \epsilon$.

(2) Für $a \geq 1$ setzen wir $x_n := \sqrt[n]{a} - 1$. Mit der Bernoulli-Ungleichung gilt $a = (1 + x_n)^n \geq 1 + nx_n$ und damit $x_n < \frac{a}{n}$. Für $\epsilon > 0$ setzen wir $N = \frac{a}{\epsilon}$ und erhalten $|\sqrt[n]{a} - 1| = x_n < \frac{a}{n} < \frac{a}{N} = \epsilon$ für alle $n > N$. Der Fall $a < 1$ folgt durch Bilden der Kehrwerte, s. Proposition 2.2.1, Fall (3).

(3) Wir setzen $x_n = \sqrt[n]{n} - 1 \geq 0$. Wir benutzen den binomischen Satz

$$(1 + x_n)^n = \sum_{i=0}^{n} \binom{n}{i} x_n^i.$$

Da $x_n \geq 0$ können wir abschätzen $n = (1 + x_n)^n \geq 1 + \binom{n}{2} x_n^2$, was äquivalent zu $n - 1 \geq \frac{n(n-1)}{2} x_n^2$ ist. Daraus folgt $x_n \leq \sqrt{\frac{2}{n}}$. Wir setzen $N = \frac{2}{\epsilon^2}$ und erhalten für alle $n > N$ dann $|\sqrt[n]{n} - 1| = x_n < \epsilon$.

(4) ist nur eine Variation zur Diskussion der geometrischen Folge, unter Benutzung des Absolutbetrags für komplexe Zahlen. □

2.2. Grenzwerte: Rechenregeln

PROPOSITION 2.2.1. *Für konvergente Folgen (a_n) und (b_n) gelten die folgenden Aussagen:*

(1) $\lim_{n\to\infty}(a_n \pm b_n) = (\lim_{n\to\infty} a_n) \pm (\lim_{n\to\infty} b_n)$
(2) $\lim_{n\to\infty}(a_n \cdot b_n) = (\lim_{n\to\infty} a_n) \cdot (\lim_{n\to\infty} b_n)$
(3) Ist $\lim_{n\to\infty} b_n \neq 0$, dann sind fast alle $b_n \neq 0$. Für ein N' so dass für alle $n > N'$ gilt $b_n \neq 0$ betrachten wir die Folge $(\frac{1}{b_n})_{n\geq N'}$. Es gilt

$$\lim_{n\to\infty} \frac{1}{b_n} = \frac{1}{b}.$$

(4) Für eine reelle Zahl c gilt $\lim_{n\to\infty}(c \cdot a_n) = c \cdot \lim_{n\to\infty} a_n$.

BEWEIS. (1) Sei $\epsilon > 0$. Dann existiert ein $N \in \mathbb{N}$ so dass für alle $n > N$ gilt $|a_n - a| < \frac{\epsilon}{2}$ und $|b_n - b| < \frac{\epsilon}{2}$. Daraus folgt $|a_n + b_n - a - b| \leq |a_n - a| + |b_n - b| < 2\frac{\epsilon}{2}$ für alle $n > N$, und damit die Behauptung.

(2) Für den Beweis benutzen wir

$$a_n b_n - ab = a_n(b_n - b) + b(a_n - a).$$

Für $\epsilon > 0$ wählen wir N, so dass für alle $n > N$ gilt

$$|a_n - a| < \min\left(\frac{\epsilon}{2|b| + 2}, 1\right), \qquad |b_n - b| < \frac{\epsilon}{2|a| + 2}.$$

Aus der ersten Ungleichung folgt $|a_n| \leq |a| + |a_n - a| < |a| + 1$. Dann können wir für alle $n > N$ abschätzen

$$|a_n b_n - ab| < \frac{(|a| + 1)\epsilon}{2|a| + 2} + \frac{|b|\epsilon}{2|b| + 2} < \epsilon.$$

(3) Wenn $\lim_{n\to\infty} b_n \neq 0$ ist, dann existiert nach Definition für $\eta = \frac{|b|}{2} > 0$ ein N', so dass für alle $n > N'$ gilt $|b_n - b| < \eta$. Damit ist dann $|b_n| \geq |b| - \eta \geq \frac{|b|}{2} > 0$. Also sind fast alle $b_n \neq 0$. Für $\epsilon > 0$ wählen wir $N > N'$, so dass für alle $n > N$ gilt $|b - b_n| < \frac{\epsilon|b|^2}{2}$. Dies bedeutet

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \frac{|b - b_n|}{|b_n||b|} < \frac{\epsilon|b|^2}{2|b_n||b|} \leq \epsilon.$$

□

ÜBUNGSAUFGABE 2.2.2. *Sei (a_n) eine Folge komplexer Zahlen mit $\lim_{n\to\infty} a_n = a$. Dann gelten*

(1) $\lim_{n\to\infty} \|a_n\| = \|a\|$
(2) $\lim_{n\to\infty} \overline{a_n} = \overline{a}$
(3) $\lim_{n\to\infty} \operatorname{Re} a_n = \operatorname{Re} a$, $\lim_{n\to\infty} \operatorname{Im} a_n = \operatorname{Im} a$
(4) $\lim_{n\to\infty} a_n = \lim \operatorname{Re} a_n + \mathrm{i} \lim \operatorname{Im} a_n$

PROPOSITION 2.2.3. *(1) Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergente Folgen. Wenn für fast alle $n \in \mathbb{N}$ gilt $a_n \leq b_n$, dann ist $\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n$.*
(2) Seien (a_n), (b_n) und (c_n) Folgen mit $a_n \leq b_n \leq c_n$ für alle $n \in \mathbb{N}$. Wenn $\lim_{n\to\infty} a_n = a = \lim_{n\to\infty} c_n$, dann gilt auch $\lim_{n\to\infty} b_n = a$.

BEWEIS. (1) Wir beweisen die Aussage mit einem Widerspruchsbeweis. Nehmen wir an, dass $a > b$ für $a = \lim_{n\to\infty} a_n$ und $b = \lim_{n\to\infty} b_n$. Für $\epsilon = \frac{a-b}{3}$ existiert nach Voraussetzung ein $N \in \mathbb{N}$, so dass for alle $n > N$ gilt $|b_n - b| < \epsilon$. Damit ist dann $|a_n - a| \geq 2\epsilon$ für alle $n > N$, also konvergiert (a_n) nicht gegen a.

(2) Sei $\epsilon > 0$. Dann existiert nach Voraussetzung ein $N \in \mathbb{N}$, so dass für alle $n > N$ gilt $|a_n - a| < \epsilon$ und $|c_n - a| < \epsilon$. Dann ist

$$a - \epsilon < a_n \leq b_n \leq c_n < a + \epsilon$$

Damit ist $|b_n - a| < \epsilon$ für alle $n > N$, und wir erhalten die Behauptung. $\square$

KOROLLAR 2.2.4. *Sei (a_n) eine konvergente Folge und $[c, d]$ ein abgeschlossenes Intervall. Wenn für fast alle Folgenglieder gilt $a_n \in [c, d]$, dann auch $\lim_{n\to\infty} a_n \in [c, d]$.*

ÜBUNGSAUFGABE 2.2.5. *Zeigen Sie, dass die folgenden Folgen konvergieren und bestimmen Sie die Grenzwerte:*
(a) $a_n = \frac{2n+3}{4n}$
(b) $b_n = \frac{n^3}{2^n}$
(c) $c_n = \frac{3^n}{n!}$
(d) $d_n = \sqrt{n}(\sqrt{n+1} - \sqrt{n})$
(e) $e_n = \sqrt[n]{\frac{1}{3} - \frac{1}{3^n}}$.

ÜBUNGSAUFGABE 2.2.6. *Die Koch-Schneeflocke entsteht aus einem gleichseitigen Dreieck der Seitenlänge a, indem im n-ten Schritt für jede Kante der Schneeflocke das mittlere Drittel entfernt und durch zwei Kanten ersetzt wird, die mit der entfernten Kante ein gleichseitiges Dreieck bilden.*
Berechnen Sie Umfang und Fläche der Koch'schen Schneeflocke.

2.3. Konvergenzkriterien

DEFINITION 2.3.1. *Eine Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen heißt*
beschränkt: *wenn es eine Zahl s gibt, so dass für alle Folgeglieder $|a_n| \leq s$ gilt.*
(streng) monoton wachsend: *wenn für alle $n \in \mathbb{N}$ gilt $a_n \leq a_{n+1}$ (bzw. $a_n < a_{n+1}$)*
(streng) monoton fallend: *wenn für alle $n \in \mathbb{N}$ gilt $a_n \geq a_{n+1}$ (bzw. $a_n > a_{n+1}$)*
monoton: *wenn sie monoton wachsend oder monoton fallend ist.*

ÜBUNGSAUFGABE 2.3.2. *Welche der Folgen sind monoton wachsend/fallend bzw. beschränkt?*
(a) $a_n = n + \frac{1}{n}$
(b) $b_n = (-1)^{n(n+1)}$
(c) $c_n = \frac{\sin(n)}{n}$
(d) $d_n = \sqrt{n-1} - \sqrt{n}$
Hinweis zu d: *Erweitern Sie mit $\sqrt{n-1} + \sqrt{n}$.*

SATZ 2.3.3. *(1) Jede konvergente Folge ist beschränkt.*
(2) Eine beschränkte und monoton wachsende Folge $(a_n)_{n\in\mathbb{N}}$ konvergiert gegen $\sup\{a_n \mid n \in \mathbb{N}\}$.
(3) Eine beschränkte und monoton fallende Folge $(a_n)_{n\in\mathbb{N}}$ konvergiert gegen $\inf\{a_n \mid n \in \mathbb{N}\}$.

BEWEIS. (1) Sei (a_n) eine konvergente Folge mit Grenzwert $a = \lim_{n\to\infty} a_n$. Nach Definition gibt es (für $\epsilon = 1$) ein N, so dass für $n > N$ gilt $|a - a_n| < 1$, d.h. $|a_n| < |a| + 1$. Dann gilt für alle n auch $|a_n| \leq \max\{|a_1|, \dots, |a_N|, |a| + 1\}$, also ist die Folge beschränkt.

(2) Wir wollen $\lim_{n\to\infty} a_n = \sup\{a_n\}$ zeigen. Das Supremum $a := \sup\{a_n\}$ auf der rechten Seite existiert, da $\{a_n \mid n \in \mathbb{N}\}$ nach Voraussetzung beschränkt (und nicht-leer) ist. Für $\epsilon > 0$ ist dann $a - \epsilon$ keine obere Schranke, also existiert ein N, so dass $a_N > a - \epsilon$. Da (a_n) monoton wachsend ist, gilt $a \geq a_n \geq a_N > a - \epsilon$ für alle $n > N$. Das bedeutet $|a_n - a| < \epsilon$, und wir haben die Konvergenz gezeigt.

(3) wird analog zu (2) bewiesen. $\square$

BEMERKUNG 2.3.4. *Die Folge* (a_n) *mit* $a_n = (-1)^n$ *ist beschränkt aber nicht konvergent.*

ÜBUNGSAUFGABE 2.3.5. *Untersuchen Sie die Folgen auf Konvergenz:*

$$a_n = \frac{1}{4^n(n+1)} \cdot \frac{(2n)!}{n!n!}, \qquad b_n = (-1)^{n+1} + \frac{1}{n}$$

DEFINITION 2.3.6. *Sei* (a_n) *eine Folge komplexer Zahlen. Eine komplexe Zahl* h *heißt* Häufungswert *der Folge* a_n*, wenn in jeder* ϵ*-Umgebung von* h *unendlich viele Folgenglieder liegen, d.h. für jedes* ϵ *unendlich viele* n *mit* $\|h - a_n\| < \epsilon$ *existieren.*

BEISPIEL 2.3.7.

- *Für eine konvergente Folge ist der Grenzwert der einzige Häufungswert.*
- *Die Folge* $a_n = \mathrm{i}^n$ *hat die vier Häufungswerte* $1, \mathrm{i}, -1, -\mathrm{i}$.
- *Eine surjektive Folge* $f\colon \mathbb{N} \to \mathbb{Q}$ *hat jede reelle Zahl als Häufungswert, da jedes Intervall unendlich viele rationale Zahlen enthählt. Solche Folgen erhalten wir zum Beispiel aus dem Cantorschen Diagonalisierungsargument.*

□

SATZ 2.3.8 (Bolzano–Weierstraß I). *Jede beschränkte Folge reeller Zahlen* (a_n) *besitzt einen größten und einen kleinsten Häufungswert. Diese werden als Limes superior* $\limsup a_n$ *bzw. Limes inferior* $\liminf a_n$ *bezeichnet. Für jedes* $\epsilon > 0$ *gilt für fast alle* n

$$a_n < \limsup a_n + \epsilon, \qquad a_n > \liminf a_n - \epsilon$$

BEWEIS. Wir beweisen die Aussagen für $\limsup$, die $\liminf$-Aussagen werden analog bewiesen.

Für die Existenz von $\limsup$ konstruieren wir rekursiv eine Intervallschachtelung $[A_k, B_k]$, so dass gilt

(1) $a_n \in [A_k, B_k]$ für unendlich viele n, und
(2) $a_n \leq B_k$ für fast alle n.

Der Rekursionsanfang ist ein Intervall $[A_1, B_1]$, das alle a_n enthält. (Dieses existiert, da (a_n) beschränkt ist.) Für den Rekursionsschritt halbieren wir das gegebene Intervall $[A_k, B_k]$ durch den Mittelpunkt $m = \frac{A_k+B_k}{2}$ und setzen

$$[A_{k+1}, B_{k+1}] := \begin{cases} [A_k, m] & \text{falls } a_n \leq m \text{ für fast alle } n \\ [m, B_k] & \text{sonst} \end{cases}$$

Dies definiert eine Intervallschachtelung, nach Konstruktion die geforderten Bedingungen erfüllt. Sei h die in allen Intervallen liegende Zahl (Intervallschachtelungsprinzip).

Wir zeigen, dass h ein Häufungswert ist. Für $\epsilon > 0$ wählen wir k, so dass $[A_k, B_k] \subset (h - \epsilon, h + \epsilon)$. Nach Konstruktion enthält $[A_k, B_k]$ unendlich viele Folgenglieder a_n, also ist h ein Häufungswert.

Die Behauptung $a_n < \limsup a_n + \epsilon$ folgt, weil nach Konstruktion $a_n \leq B_k < h + \epsilon$ für fast alle n gilt.

Es bleibt noch zu zeigen, dass kein $h' > h$ ein Häufungswert sein kann. Das folgt aus $a_n < \limsup a_n + \epsilon$. Wir setzen $\epsilon_0 = \frac{h'-h}{2}$ und haben dann $a_n < h + \epsilon_0 = h' - \epsilon_0$ für fast alle n. Damit liegen nur endlich viele a_n in der offenen ϵ_0-Umgebung von h', also kann h' kein Häufungswert sein. □

ÜBUNGSAUFGABE 2.3.9. *Seien* (a_n) *und* (b_n) *beschränkte Folgen mit nichtnegativen Gliedern. Zeigen Sie, dass* $\limsup_n (a_b \cdot b_b) \leq (\limsup_n a_n) \cdot (\limsup_n b_n)$*. Wenn zusätzlich* (a_n) *konvergent ist, dann gilt* $\limsup_n (a_n b_n) = (\lim_n a_n)(\limsup_n b_n)$*.*

DEFINITION 2.3.10. *Ist (a_n) eine Folge und $(n_k)_{k\in\mathbb{N}}$, $n_k \in \mathbb{N}$, eine streng monoton wachsende Folge von Indizes, dann heißt die Folge $(a_{n_k})_{k\in\mathbb{N}}$ eine* Teilfolge von (a_n).

LEMMA 2.3.11. *Eine Zahl h ist genau dann Häufungswert einer Folge (a_n), wenn h der Grenzwert einer Teilfolge (a_{n_k}) ist.*

BEWEIS. Sei h der Grenzwert einer konvergenten Teilfolge (a_{n_k}). Dann enthält jede ϵ-Umgebung $U_\epsilon(h)$ von h fast alle a_{n_k}, also unendlich viele a_n. Damit ist h ein Häufungswert von (a_n).

Sei h ein Häufungswert von (a_n). Das bedeutet, dass jede ϵ-Umgebung $U_\epsilon(h)$ unendlich viele a_n enthält. Wir konstruieren rekursiv eine konvergente Teilfolge. Für den Rekursionsanfang wählen wir ein a_{n_1}, das in $U_1(h)$ liegt. Haben wir im k-ten Schritt a_{n_k}, wählen wir einen Index $n_{k+1} > n_k$, so dass $a_{n_{k+1}} \in U_{\frac{1}{k+1}}(h)$. Nach Konstruktion ist $|a_{n_k} - h| < \frac{1}{k}$, also gilt $h = \lim_{k\to\infty} a_{n_k}$. □

SATZ 2.3.12 (Bolzano–Weierstraß II). *Jede beschränkte Folge reeller oder komplexer Zahlen besitzt eine konvergente Teilfolge.*

BEWEIS. Die Aussage für reelle Zahlen folgt aus Satz 2.3.8 und Lemma 2.3.11. Für eine Folge $(z_n = a_n + \mathrm{i}b_n)_{n\in\mathbb{N}}$ von komplexe Zahlen betrachten wir die Folgen (a_n) und (b_n) der Real- bzw. Imaginärteile. Aus (a_n) wählen wir eine konvergente Teilfolge mit Indizes $(n_k)_{k\in\mathbb{N}}$ aus, dann wählen wir aus $(b_{n_k})_{k\in\mathbb{N}}$ ebenfalls wieder eine konvergente Teilfolge aus. Damit haben wir eine konvergente Teilfolge in (z_n) gefunden. □

KOROLLAR 2.3.13 (Bolzano–Weierstraß I für komplexe Folgen). *Jede beschränkte Folge komplexer Zahlen besitzt einen Häufungswert.*

BEWEIS. Folgt aus Bolzano–Weierstraß II, Satz 2.3.12: der Grenzwert einer konvergenten Teilfolge ist ein Häufungswert. □

SATZ 2.3.14 (Cauchy-Kriterium). *Eine Folge (a_n) konvergiert genau dann, wenn es zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass $|a_n - a_m| < \epsilon$ für alle $n, m > N$ gilt.*

BEWEIS. $\Rightarrow$ Wir nehmen an, dass die Folge konvergiert mit $a = \lim_{n\to\infty} a_n$. Dann existiert zu jedem $\epsilon > 0$ ein N, so dass $|a_k - a| < \frac{\epsilon}{2}$ für alle $k > N$ ist. Dann haben wir für alle $n, m > N$

$$|a_n - a_m| \leq |a_n - a| + |a - a_m| < \epsilon$$

$\Leftarrow$ Nehmen wir an, dass die Bedingung erfüllt ist.

Wir zeigen zuerst, dass die Folge beschränkt ist: Nach Voraussetzung existiert ein N, so dass $|a_n - a_m| < 1$ für alle $n, m > N$ gilt. Daraus folgt $|a_n| < |a_{N+1}| + 1$ für $n > N$. Die Folge ist damit durch $\max\{|a_1|, \ldots, |a_N|, |a_{N+1}| + 1\}$ beschränkt.

Nach dem Satz von Bolzano–Weierstraß 2.3.12 existiert eine konvergente Teilfolge (a_{n_k}) mit Grenzwert $a = \lim_{k\to\infty} a_{n_k}$. Wir zeigen, dass $\lim_{n\to\infty} a_n = a$. Für $\epsilon > 0$ existiert ein N', so dass $|a_n - a_m| < \frac{\epsilon}{2}$ für alle $n, m > N'$ gilt. Ebenfalls existiert ein $n_k > N'$, so dass $|a_{n_k} - a| < \frac{\epsilon}{2}$. Damit haben wir für alle $n > N'$

$$|a_n - a| \leq |a_n - a_{n_k}| + |a_{n_k} - a| < \epsilon.$$

Damit konvergiert die Folge (a_n) gegen a. □

DEFINITION 2.3.15. *Eine Folge (a_n) komplexer Zahlen heißt* Cauchy-Folge*, wenn zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass für alle $n, m > N$ gilt*

$$\|a_n - a_m\| < \epsilon.$$

ÜBUNGSAUFGABE 2.3.16. *Zeigen Sie, dass $x_n = \frac{(-1)^n}{n^2}$ eine Cauchy-Folge ist.*

SATZ 2.3.17. *Für einen archimedisch angeordneten Körper sind die folgenden Aussagen sind äquivalent:*

(1) *Supremumseigenschaft*
(2) *Intervallschachtelungsprinzip*
(3) *Satz von Bolzano–Weierstraß*
(4) *Jede Cauchy-Folge konvergiert.*

BEWEIS. Die Äquivalenz von Supremumseigenschaft und Intervallschachtelungsprinzip haben wir in Satz 1.6.24 bewiesen. Wir müssen also nur noch die Implikationen (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (2) zeigen. Die Implikation (2) $\Rightarrow$ (3) ist Satz 2.3.8, die Implikation (3) $\Rightarrow$ (4) folgt aus Satz 2.3.14.

Es bleibt noch die Implikation (4) $\Rightarrow$ (2). Sei $(I_n)_{n\in\mathbb{N}}$ mit $I_n = [a_n, b_n]$ eine Intervallschachtelung. Wir zeigen, dass $(a_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge ist. Für $\epsilon > 0$ existiert N mit $b_N - a_N < \epsilon$ und aus der Intervallschachtelungsbedingung $I_{n+1} \subseteq I_n$ folgt $a_n, a_m \in [a_N, b_N]$ für alle $n, m > N$. Das bedeutet $|a_m - a_n| < \epsilon$, wir haben also eine Cauchy-Folge. Nach (4) existiert ein Grenzwert $s = \lim a_n$ und wir wollen zeigen, dass $s \in [a_n, b_n]$ für alle n gilt. Die unteren Intervallgrenzen (a_n) wachsen monoton, also ist $a_n \leq s$ für alle n. Aus $a_k \leq b_n$ für alle k, n folgt $s \leq b_n$ für alle n. Also ist $s \in [a_n, b_n]$ für alle n, und das Intervallschachtelungsprinzip folgt. □

ÜBUNGSAUFGABE 2.3.18. *Zeigen Sie die folgenden Aussagen:*

(1) *Jede reelle Zahl ist Grenzwert einer monotonen beschränkten Folge von rationalen Zahlen.*
(2) *Jede Folge reeller Zahlen besitzt eine monotone Teilfolge.*
(3) *Wenn eine beschränkte Folge reeller Zahlen genau einen Häufungswert hat, dann konvergiert sie gegen den Häufungswert.*

ÜBUNGSAUFGABE 2.3.19. *(Teilfolgen) Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge. Wir betrachten die beiden Teilfolgen $b_n = a_{2n}$ und $c_n = a_{2n+1}$. Zeigen Sie die folgenden Aussagen:*

(a) *Konvergieren (b_n) und (c_n) beide gegen den Grenzwert a, dann konvergiert auch (a_n) gegen a.*
(b) *Ist (a_n) konvergent, dann sind auch (b_n) und (c_n) konvergent.*
(c) *Es gibt divergente Folgen (a_n), so dass (b_n) und (c_n) konvergieren.*

ÜBUNGSAUFGABE 2.3.20. *Seien (a_n) und (b_n) beschränkte Folgen mit nichtnegativen Gliedern. Zeigen Sie, dass $\limsup_n (a_b \cdot b_b) \leq (\limsup_n a_n) \cdot (\limsup_n b_n)$. Wenn zusätzlich (a_n) konvergent ist, dann gilt $\limsup_n (a_n b_n) = (\lim_n a_n)(\limsup_n b_n)$.*

2.4. Anwendungen

Grenzwerte von Folgen sind eine sehr allgemeine Methode zur Definition von interessanten reellen Zahlen, sowie zur Approximation bzw. näherungsweisen Berechnung von solchen Zahlen. Wir diskutieren ein paar Beispiele.

2.4.1. Allgemeine Potenz.

ÜBUNGSAUFGABE 2.4.1. *Sei r_n eine Folge rationaler Zahlen mit $\lim_{n\to\infty} r_n = r \in \mathbb{Q}$. Dann gilt für alle $a > 0$ auch $\lim_{n\to\infty} a^{r_n} = a^r$.*

Hinweis: *Führen Sie die Aussage auf den Spezialfall $r = 0$ zurück und beweisen Sie diesen.*

DEFINITION 2.4.2. *Seien a und ρ zwei reelle Zahlen. Um die allgemeine Potenz a^ρ zu definieren, wählen wir eine monoton wachsende Folge rationaler Zahlen (r_n) mit $\lim_{n\to\infty} r_n = \rho$. Die Folge $(a^{r_n})_{n\in\mathbb{N}}$ ist monoton wachsend und beschränkt, also konvergent, und wir definieren $a^\rho := \lim_{n\to\infty} a^{r_n}$.*

ÜBUNGSAUFGABE 2.4.3. *Seien a und ρ zwei reelle Zahlen. Zeigen Sie, dass die Definition von a^ρ in Definition 2.4.2 nicht von der gewählten Folge (r_n) mit $\lim_{n\to\infty} r_n = \rho$ abhängt. Allgemeiner gilt für jede Folge (s_n) rationaler Zahlen (nicht notwendig monoton) mit $\lim_{n\to\infty} s_n = \rho$ auch $a^\rho = \lim_{n\to\infty} a^{s_n}$.*

Hinweis: *Beutzen Sie Übungsaufgabe 2.4.1.*

ÜBUNGSAUFGABE 2.4.4. *(a) Zeigen Sie, dass di allgemeine Potenz aus Definition 2.4.2 die Potenzgesetze erfüllt:*

$$a^r a^s = a^{r+s}, \quad (a^r)^s = a^{rs}, \quad a^r b^r = (ab)^r.$$

(b) Zeigen Sie, dass für reelle Zahlen $a < b$ und $r > 0$ gilt $a^r < b^r$.

2.4.2. Approximation von Quadratwurzeln. Für $a > 0$ betrachten wir die Folge (x_n), die für einen beliebigen Startwert $x_0 > 0$ durch die Rekursion

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$$

gegeben ist.

PROPOSITION 2.4.5. *Die Folge $(x_n)_{n\in\mathbb{N}}$ konvergiert für beliebig gewählte Startwerte $x_0 > 0$ gegen $\sqrt{a}$.*

BEWEIS. Mit Induktion zeigen wir $x_n > 0$. Der Induktionsanfang ist die Voraussetzung $x_0 > 0$ an den Startwert, im Induktionsschritt folgt $x_{n+1} > 0$ mit der rekursiven Definition aus $x_n > 0$. Die Folge ist also wohldefiniert (wir teilen nicht durch 0). Wir haben

$$x_{n+1}^2 - a = \frac{1}{4}\left(x_n^2 + 2a + \frac{a^2}{x_n^2}\right) - a = \frac{1}{4}\left(x_n - \frac{a}{x_n}\right)^2 \geq 0,$$

also gilt $x_n \geq \sqrt{a}$ für alle $n \geq 1$, insbesondere ist (x_n) nach unten beschränkt. Aus $x_n \geq \sqrt{a}$ für $n \geq 0$ folgt mit

$$x_{n+1} - x_n = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) - x_n = -\frac{x_n}{2} + \frac{a}{2x_n} \leq 0,$$

dass die Folge (x_n) ab $n = 1$ monoton fallend ist. Also konvergiert (x_n). Um den Grenzwert $l = \lim_{n\to\infty} x_n$ zu ermitteln, bilden wir den Grenzwert $n \to \infty$ der Rekursionsvorschrift und erhalten die Gleichung

$$l = \frac{1}{2}\left(l + \frac{a}{l}\right).$$

Umstellen liefert $l^2 = a$, was die Behauptung war. □

ÜBUNGSAUFGABE 2.4.6. *Eine alternative Folge für die Berechnung der Quadratwurzeln ist die Folge $\left(\frac{a_n}{b_n}\right)_{n\in\mathbb{N}}$ mit den beiden rekursiv definierten Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$:*

$$a_1 = 1, \qquad b_1 = 1, \qquad a_n = a_{n-1} + 2b_{n-1}, \qquad b_n = a_{n-1} + b_{n-1}.$$

Zeigen Sie, dass die Folge $\left(\frac{a_n}{b_n}\right)_{n\in\mathbb{N}}$ gegen $\sqrt{2}$ konvergiert.

Hinweis: *Eine Variante ist die folgende Vorgehensweise: Zeigen Sie zuerst induktiv, dass die Folgenglieder (a_n) und (b_n) die Gleichung $a^2 - 2b^2 = \pm 1$ lösen. Daraus folgt, dass die Folge $\left(\frac{a_n}{b_n}\right)$ beschränkt ist. Dann können wir die beiden Teilfolgen $\frac{a_{2n+1}}{b_{2n+1}}$ und $\frac{a_{2n}}{b_{2n}}$ betrachten. Induktiv beweisen wir, dass die erste Teilfolge monoton wachsend, die zweite monoton fallend ist. Also konvergieren beide Teilfolgen, und zwar gegen den gemeinsamen Grenzwert $\sqrt{2}$.*

2.4.3. Kettenbrüche.

ÜBUNGSAUFGABE 2.4.7. *Die positive Zahl g, welche $g = 1 + \frac{1}{g}$ erfüllt, heißt* goldener Schnitt.

(a) *Bestimmen Sie g.*
(b) *Es sei $(x_n)_n$ die rekursiv durch $x_0 = 1$ und $x_{n+1} = 1 + \frac{1}{x_n}$ definierte Folge. Zeigen Sie mit vollständiger Induktion, dass*

$$|x_n - g| \leq \frac{1}{g^{n+1}}$$

gilt.
(c) *Zeigen Sie, dass $(x_n)_n$ gegen g konvergiert.*
(d) *Überlegen Sie sich, dass Aussage (c) als die folgende Gleichung umformuliert werden kann:*

$$1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cdots}}}} = g$$

ÜBUNGSAUFGABE 2.4.8. *Beweisen Sie die folgende Aussage:*

$$\sqrt{a + \sqrt{a + \sqrt{a + \cdots}}} = \frac{1 + \sqrt{1 + 4a}}{2}.$$

Betrachten Sie dafür die rekursive Folge $x_0 = \sqrt{a}$ und $x_{n+1} = \sqrt{a + x_n}$ und zeigen Sie die folgenden Aussagen:

(a) *Zeigen Sie, dass diese Folge monoton wachsend ist.*
(b) *Zeigen Sie, dass die Folge $(x_n)_n$ durch $1 + \sqrt{a}$ nach oben beschränkt ist.*
(c) *Schließen Sie aus der Konvergenz den behaupteten Grenzwert.*

2.4.4. Wachstumsverhalten von Folgen. Wir können auch durch geeignete Konstruktionen das Wachstumsverhalten von Folgen vergleichen:

DEFINITION 2.4.9. *Zwei Folgen (a_n) und (b_n) (mit $b_n \neq 0$ für alle $n \in \mathbb{N}$) heißen* asymptotisch gleich, *wenn*

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 1$$

ÜBUNGSAUFGABE 2.4.10. *Sei $f(x) = \sum_{i=0}^{d} a_i x^i$ ein Polynom mit Leitkoeffizient $a_d \neq 0$. Dann sind die Folgen $(f(n))_{n\in\mathbb{N}}$ und $(a_d n^d)_{n\in\mathbb{N}}$ asymptotisch gleich.*

Andererseits können wir sagen, dass die Folge (b_n) deutlich schneller wächst, als die Folge (a_n), wenn

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 0.$$

Dies wird als $(a_n) \prec (b_n)$ notiert.

BEISPIEL 2.4.11. *Für natürliche Zahlen $k < l$ wächst $(n^k)_{n\in\mathbb{N}}$ deutlich langsamer als $(n^l)_{n\in\mathbb{N}}$. Ebenso wächst $(\sqrt{n})$ deutlich langsamer als (n).* □

BEISPIEL 2.4.12. *Für eine komplexe Zahl z mit $\|z\| > 1$ wachsen die Potenzen von z deutlich schneller als Polynome in den Exponenten. (Exponentielles Wachstum ist deutlich schneller als polynomiales Wachstum.) Genauer haben wir für alle $k \in \mathbb{N}$ und alle $z \in \mathbb{C}$, $\|z\| > 1$:*

$$\lim_{n\to\infty} \frac{n^k}{z^n} = 0$$

Dafür setzen wir $\|z\| = 1 + x$ *mit* $x > 0$*. Für* $n > 2k$ *folgt aus dem binomischen Satz und* $x > 0$

$$\begin{aligned}(1+x)^n &= \sum_{i=0}^{n} \binom{n}{i} x^i \\ &> \binom{n}{k+1} x^{k+1} = \frac{n(n-1)\cdots(n-k)}{(k+1)!} x^{k+1} \\ &> \frac{n^{k+1}}{2^{k+1}(k+1)!} x^{k+1}\end{aligned}$$

Dabei nutzen wir im letzten Schritt $n - k > \frac{n}{2}$*. Umstellen liefert*

$$\left\| \frac{n^k}{z^n} \right\| < \frac{2^{k+1}(k+1)!}{x^{k+1}} \cdot \frac{1}{n}$$

Für $\epsilon > 0$ *ist dann* $\left\| \frac{n^k}{z^n} \right\| < \epsilon$ *für alle*

$$n > \max\left(\frac{2^{k+1}(k+1)!}{x^{k+1}} \cdot \frac{1}{\epsilon}, 2k \right).$$

□

BEMERKUNG 2.4.13. *Solche Aussagen zum qualitativen Vergleich des Wachstums von Folgen sind zum Beispiel für Komplexitätsabschätzungen von Algorithmen relevant. Die Folgenglieder sind dabei z.B. die Anzahl an Rechenschritten, in Abhängigkeit von der Größe* n *der Eingabe. Dafür wird auch die* Bachmann–Landau-Notation *verwendet. Als Beispiel haben wir die folgenden Aussagen zum Folgenwachstum:*

$$\log_b n \prec n^r \prec a^n \prec n! \prec n^n$$

Logarithmisches Wachstum (z.B. bei der Binärsuche in einem sortierten Array) ist langsamer als polynomiales Wachstum (einfache Suchalgorithmen quadratisch oder Matrizenmultiplikation kubisch), polynomiales Wachstum langsamer als exponentielles Wachstum (travelling salesman mit dynamischer Programmierung), exponentielles Wachstum langsamer als faktorielles Wachstum (brute force travelling salesman).

ÜBUNGSAUFGABE 2.4.14. *Zeigen Sie*

(a) *Für alle reellen* $a > 0$ *gilt*

$$\lim_{n\to\infty} \frac{a^n}{n!} = 0$$

(b)

$$\lim_{n\to\infty} \frac{n!}{n^n} = 0$$

2.4.5. Weitere Aufgaben.

ÜBUNGSAUFGABE 2.4.15. *Für* $a > 0$ *definieren wir eine Folge durch* $x_0 \in (0, \frac{2}{a})$ *beliebig und* $x_{n+1} := x_n(2 - ax_n)$*. Zeigen Sie, dass die Folge ab* $n = 1$ *monoton wachsend ist und gegen* $\frac{1}{a}$ *konvergiert.*

ÜBUNGSAUFGABE 2.4.16. *Zeigen Sie, dass die Folge* $a_n = \left(1 + \frac{1}{n}\right)^n$ *monoton wachsend und durch* 3 *nach oben beschränkt ist. Der Grenzwert ist die Eulersche Zahl* e.

Hinweis: *Benutzen Sie den binomischen Satz, geometrische Summe und die Abschätzung*

$$\frac{1}{m^k}\binom{m}{k} < \frac{1}{n^k}\binom{n}{k} \le \frac{1}{k!} \le \frac{1}{2^{k-1}}$$

für natürliche Zahlen $n > m$ *und* $k = 2, \ldots, n$.

2.5. Unendliche Reihen, Konvergenz und Beispiele

DEFINITION 2.5.1.

- *Sei $(a_k)_{k\in\mathbb{N}}$ eine Folge. Durch $S_n = \sum_{k=0}^n a_k$ wird eine Folge $(S_n)_{n\in\mathbb{N}}$ definiert, die zu $(a_k)_{k\in\mathbb{N}}$ gehörende* unendliche Reihe.
- *Die S_n heißen* Partialsummen. *Wir sagen, dass* die Reihe konvergiert, *wenn die Folge $(S_n)_{n\in\mathbb{N}}$ konvergiert. Im Konvergenzfall heißt*

$$\sum_{k=0}^{\infty} a_k := \lim_{n\to\infty} S_n = \lim_{n\to\infty} \sum_{k=0}^{n} a_k.$$

die Summe *der Reihe.*
- *Eine Reihe $\sum_{k=0}^{\infty} a_k$ heißt* absolut konvergent, *wenn die Reihe $\sum_{k=0}^{\infty} |a_k|$ konvergent ist.*

BEMERKUNG 2.5.2. *Eine Reihe heißt* divergent, *wenn sie nicht konvergent ist. Konvergenz von Reihen ist nur ein Spezialfall der Konvergenz von Folgen, angewendet auf die Folge der Partialsummen. Dennoch gibt es für Reihen ein paar spezifischere Konvergenzkriterien, diese benutzen die zusätzliche Struktur in der Partialsummenfolge.*

ÜBUNGSAUFGABE 2.5.3. *Sei $\sum_{n=0}^{\infty} z_n$ eine komplexe Reihe. Zeigen Sie, dass die Reihe genau dann konvergiert, wenn die Reihen $\sum_{n=0}^{\infty} \mathrm{Re}(z_n)$ und $\sum_{n=0}^{\infty} \mathrm{Im}(z_n)$ konvergieren, und dass für die Summe gilt*

$$\sum_{n=0}^{\infty} z_n = \sum_{n=0}^{\infty} \mathrm{Re}(z_n) + \mathrm{i} \sum_{n=0}^{\infty} \mathrm{Im}(z_n)$$

BEISPIEL 2.5.4. *Wir nennen $\sum_{i=0}^{\infty} x^i$ die* geometrische Reihe. *Die Partialsummen sind nach Proposition 1.3.14*

$$S_n = \sum_{i=0}^{n} x^i = \begin{cases} \frac{1-x^{n+1}}{1-x} & x \neq 1 \\ n+1 & x = 1. \end{cases}$$

Das bedeutet, dass die geometrische Reihe für $|x| \geq 1$ nicht konvergiert. Für $|x| < 1$ konvergiert die geometrische Reihe und wir haben (mit $\lim_{n\to\infty} x^n = 0$)

$$\sum_{i=0}^{\infty} x^i = \frac{1}{1-x}.$$

Für eine Veranschaulichung der geometrischen Reihe für $x = \frac{1}{2}$, s. Abbildung 6.

ABBILDUNG 6. Veranschaulichung der geometrischen Reihe mit $x = \frac{1}{2}$

- *Im Paradox von Zeno von Elea wird ein Wettlauf zwischen Achilles und einer Schildkröte betrachtet. Achilles läuft doppelt so schnell wie die Schildkröte, die deswegen mit 10 Meter Vorsprung starten darf. Wenn Achilles die 10 Meter bis zum Startpunkt der Schildkröte zurückgelegt hat, ist die Schildkröte 5 Meter weiter. Hat Achilles diese 5 Meter geschafft, ist die*

Schildkröte immer noch 2,5 Meter entfernt. Und so weiter. Zeno ging davon aus, dass diese unendliche Summation dazu führt, dass Achilles die Schildkröte nie einholen kann, was paradox ist (da es der Alltagserfahrung mit Schildkröten widerspricht). Tatsächlich haben wir hier aber eine geometrische Reihe $\sum_{i=0}^{\infty} \frac{1}{2^i} = 2$*, deren Wert endlich ist: Nach dem doppelten Vorsprung hat Achilles die Schildkröte eingeholt.*

- *Eine Tafel Schokolade kostet 1 Euro. In jeder Tafel ist ein Bonuspunkt, für 10 Bonuspunkte gibt es eine Tafel Schokolade. Wie viel ist eine Tafel Schokolade (im Idealfall) tatsächlich wert?*

□

ÜBUNGSAUFGABE 2.5.5. *(a) Wie hoch war die durchschnittliche jährliche Inflationsrate, wenn sich der Preis für den Standardwarenkorb innerhalb von 15 Jahren von 100 auf 150 Euro erhöht hat?*

(b) Wir würfeln so lange, bis eine 1 erscheint. Wie groß ist die Wahrscheinlichkeit, dass vor der 1 keine gerade Zahl gewürfelt wird? (Standardwürfel, sechs Seiten, alle gleiche Wahrscheinlichkeit, kein Gedächtnis)

BEISPIEL 2.5.6. *Wir nennen* $\sum_{k=1}^{\infty} \frac{1}{k}$ *die* harmonische Reihe. *Zur Konvergenzbetrachtung schätzen wir die Partialsummen wie folgt ab:*

$$1 + \frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{\geq 2 \cdot \frac{1}{4}} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{\geq 4 \cdot \frac{1}{8}} + \cdots$$

Dann folgt für alle i

$$S_{2^i} \geq 1 + i \cdot \frac{1}{2}.$$

Die Folge (S_k) *wächst also unbeschränkt, also konvergiert* $\sum_{k=1}^{\infty} \frac{1}{k}$ *nicht.* □

BEMERKUNG 2.5.7. *In Festkomma-Arithmetik "konvergiert" die harmonische Reihe (das ist allgemeiner der Fall für Reihen, deren Koeffizientenfolgen Nullfolgen sind – ab einem bestimmten Punkt sind alle weiteren aufzusummierenden Folgeglieder betragsmäßig kleiner als die Genauigkeit der Festkomma-Arithmetik). Nach der obigen Abschätzung liegt der "Wert" der Reihe auf einem Festkomma-Arithmetik-Rechner im Bereich der halben Bitzahl.*

ÜBUNGSAUFGABE 2.5.8. *(a) Wie hoch war die durchschnittliche jährliche Inflationsrate, wenn sich der Preis für den Standardwarenkorb innerhalb von 15 Jahren von 100 auf 150 Euro erhöht hat?*

(b) Wir würfeln so lange, bis eine 1 erscheint. Wie groß ist die Wahrscheinlichkeit, dass vor der 1 keine gerade Zahl gewürfelt wird. (Standardwürfel, sechs Seiten, alle gleiche Wahrscheinlichkeit, kein Gedächtnis)

2.6. Konvergenzkriterien und Rechenregeln für Reihen

PROPOSITION 2.6.1 (Cauchy-Kriterium). *Die Reihe* $\sum_{i=0}^{\infty} a_i$ *mit Partialsummenfolge* $(S_n = \sum_{i=0}^{n} a_i)_{n\in\mathbb{N}}$ *konvergiert genau dann, wenn für alle* $\epsilon > 0$ *ein* $N \in \mathbb{N}$ *existiert, so dass für alle* $m > n > N$ *gilt* $|S_m - S_n| = \left|\sum_{i=n+1}^{m} a_i\right| < \epsilon$*.*

BEWEIS. Umformulierung von Satz 2.3.14. □

KOROLLAR 2.6.2. *Wenn die Reihe* $\sum_{i=0}^{\infty} a_i$ *konvergiert, dann ist die Folge* $(a_i)_{i\in\mathbb{N}}$ *eine Nullfolge.*

BEWEIS. Aus dem Cauchy-Kriterium für die Konvergenz der Reihe erhalten wir insbesondere, dass für alle $\epsilon > 0$ ein N existiert, so dass für alle $n > N$ gilt $|S_{n+1} - S_n| = |a_{n+1}| < \epsilon$. Das ist die Behauptung. □

BEMERKUNG 2.6.3. *Die harmonische Reihe zeigt, dass die Umkehrung dieser Aussage nicht gilt.*

Die Rechenregeln für Addition von Grenzwerten, s. Proposition 2.2.1, liefern Rechenregeln für Addition von Reihen:

KOROLLAR 2.6.4. *Seien $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ zwei konvergente Reihen. Dann sind die Reihen $\sum_{k=0}^{\infty}(a_k \pm b_k)$ und $\sum_{k=0}^{\infty} ca_k$ für $c \in \mathbb{R}$ konvergent und es gilt:*

$$\sum_{k=0}^{\infty}(a_k \pm b_k) = \left(\sum_{k=0}^{\infty} a_k\right) \pm \left(\sum_{k=0}^{\infty} b_k\right), \qquad \sum_{k=0}^{\infty} c \cdot a_k = c \cdot \sum_{k=0}^{\infty} a_k.$$

PROPOSITION 2.6.5 (Leibniz-Kriterium). *Sei $(a_i)_{i\in\mathbb{N}}$ eine monoton fallende Nullfolge. Dann konvergiert die Reihe $\sum_{i=0}^{\infty}(-1)^i a_i$.*

BEWEIS. Für die Partialsummen $S_n = \sum_{i=0}^{n}(-1)^i a_i$ betrachten wir die Folge (S_{2n}). Wegen

$$S_{2n+2} - S_{2n} = (-1)^{2n+2}(a_{2n+2} - a_{2n+1}) \leq 0$$

und $S_{2n} = (a_0 - a_1) + (a_2 - a_3) + \cdots + (a_{2n-2} - a_{2n-1}) + a_{2n} \geq a_{2n} \geq 0$ ist $(S_{2n})_{n\in\mathbb{N}}$ eine monoton fallende nach unten beschränkte Folge, damit konvergent. Analog ist $(S_{2n+1})_{n\in\mathbb{N}}$ eine monoton wachsende durch a_0 nach oben beschränkte Folge, damit konvergent. Dann ist

$$\lim_{n\to\infty} S_{2n} = \lim_{n\to\infty}(S_{2n+1} - a_{2n+1}) = \lim_{n\to\infty} S_{2n+1},$$

da (a_n) eine Nullfolge ist. Die Reihe konvergiert also gegen den gemeinsamen Grenzwert von (S_{2n}) und (S_{2n+1}). □

BEISPIEL 2.6.6. *Die Leibniz-Reihe $\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1}$ konvergiert nach dem Leibniz-Kriterium. Der Wert der Reihe ist*

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = \frac{\pi}{4},$$

s. Kapitel zu trigonometrischen Funktionen. □

PROPOSITION 2.6.7. *(1) Absolute Konvergenz impliziert Konvergenz.*
(2) Die Reihe $\sum_{k=0}^{\infty} a_k$ ist genau dann absolut konvergent, wenn die Folge $(S_n = \sum_{k=0}^{n} |a_k|)_{n\in\mathbb{N}}$ der Partialsummen beschränkt ist.

BEWEIS. (1) Wir benutzen das Cauchy-Kriterium von Proposition 2.6.1. Die Reihe $\sum_{k=0}^{\infty} a_k$ konvergiert, wenn es für jedes $\epsilon > 0$ ein N gibt, so dass für alle $m > n > N$ gilt $|\sum_{i=n+1}^{m} a_i| < \epsilon$. Mit Dreiecksungleichung ist $|\sum_{i=n+1}^{m} a_i| \leq \sum_{i=n+1}^{m} |a_i|$. Damit folgt die Konvergenzbedingung für $\sum_{k=0}^{\infty} a_k$ aus der Konvergenzbedingung für $\sum_{k=0}^{\infty} |a_k|$.

(2) Die Partialsummenfolge für $\sum_{k=0}^{\infty} |a_k|$ ist monoton wachsend. Damit konvergiert die Partialsummenfolge für $\sum_{k=0}^{\infty} |a_k|$ genau dann, wenn sie beschränkt ist, s. Satz 2.3.3. □

BEISPIEL 2.6.8. *Die alternierende harmonische Reihe $\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k}$ konvergiert nach dem Leibniz-Kriterium, aber nicht absolut nach Beispiel 2.5.6. Der Wert der Reihe ist $\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k} = \ln 2$, s. Kapitel zu Logarithmus-Funktion.* □

PROPOSITION 2.6.9 (Vergleichskriterium/Majorantenkriterium). *Seien (a_k) und (b_k) zwei Folgen mit $0 \leq |a_k| \leq b_k$ für fast alle k. Dann gilt*

(1) *Wenn* $\sum_{k=0}^{\infty} b_k$ *konvergent ist, dann ist* $\sum_{k=0}^{\infty} a_k$ *absolut konvergent. Für die Grenzwerte gilt dann*

$$\sum_{k=0}^{\infty} |a_k| \leq \sum_{k=0}^{\infty} b_k.$$

(2) *Wenn* $\sum_{k=0}^{\infty} a_k$ *divergent ist, dann ist auch* $\sum_{k=0}^{\infty} b_k$ *divergent.*

Beweis. (2) folgt wegen $(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$ aus (1).

(1) Nach Definition der Konvergenz von $\sum_{k=0}^{\infty} b_k$ existiert für jedes $\epsilon > 0$ ein N so dass für alle natürlichen $m > n > N$ gilt $\sum_{i=n+1}^{m} b_k < \epsilon$. Aus der Voraussetzung folgt, dass wir durch eventuelles Vergrößern von N für $k \geq N$ auch $0 \leq |a_k| \leq b_k$ haben. Mit der Dreiecksungleichung folgt daraus für alle $m > n > N$

$$\left|\sum_{i=n+1}^{m} a_i\right| \leq \sum_{i=n+1}^{m} |a_i| \leq \sum_{i=n+1}^{m} b_i < \epsilon$$

Die Aussage über die Grenzwerte folgt aus Proposition 2.2.3. □

Beispiel 2.6.10. *Eine Potenzreihe* $\sum_{k=0}^{\infty} a_k z^k$ *mit komplexen Koeffizienten* a_k *konvergiert für* $\|a_k\| \leq 1$ *für alle* k *und* $\|z\| < 1$ *durch Vergleich mit der geometrischen Reihe.* □

Proposition 2.6.11 (Quotientenkriterium). *Sei* (a_k) *eine Folge mit* $a_k \neq 0$ *für fast alle* k. *Wenn* $\lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| = a$, *dann gilt*

(1) *Falls* $a < 1$ *ist, ist* $\sum_{k=0}^{\infty} a_k$ *absolut konvergent.*
(2) *Falls* $a > 1$ *ist, ist* $\sum_{k=0}^{\infty} a_k$ *divergent.*

Beweis. Wir beweisen die Aussage durch Vergleich mit der geometrischen Reihe.

(1) Sei $\lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| < 1$. Dann existiert $q \in (0,1)$ und $N \in \mathbb{N}$, so dass für alle $k \geq N$ gilt

$$\left|\frac{a_{k+1}}{a_k}\right| \leq q.$$

Dann gilt auch $|a_{N+l}| \leq q^l |a_N|$. Wir betrachten die geometrische Folge $b_k = |a_N| q^{k-N}$ für $k \geq N$ und $b_k = 0$ für $k < N$. Die Reihe $\sum_{k=0}^{\infty} b_k$ konvergiert, und das Vergleichskriterium aus Proposition 2.6.9 liefert die Behauptung.

(2) Aus $\lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| > 1$ folgt, dass die Folge $(|a_k|)$ ab einem bestimmten Index streng monoton wächst, also keine Nullfolge sein kann. □

Beispiel 2.6.12. *Für* $\sum_{k=0}^{\infty} \frac{1}{k!}$ *ist* $\frac{a_{k+1}}{a_k} = \frac{1}{k+1}$, *damit konvergiert die Reihe nach dem Quotientenkriterium. Der Grenzwert ist die Eulersche Zahl* $\mathrm{e} = \sum_{k=0}^{\infty} \frac{1}{k!}$. *Allgemeiner konvergiert* $\sum_{n=0}^{\infty} \frac{z^n}{n!}$, *s. Kapitel zur Exponentialfunktion.* □

Proposition 2.6.13 (Wurzelkriterium). *Sei* (a_k) *eine Folge, so dass*

$$\limsup_{k\to\infty} \sqrt[k]{|a_k|} = a$$

existiert. Dann gilt:

(1) *Wenn* $a < 1$ *ist, dann konvergiert* $\sum_{k=0}^{\infty} a_k$ *absolut.*
(2) *Wenn* $a > 1$ *ist, dann divergiert* $\sum_{k=0}^{\infty} a_k$.

Beweis. Die Aussage wird ebenfalls wieder durch Vergleich mit der geometrischen Reihe bewiesen. Für (1) folgt aus $\limsup_{k\to\infty} \sqrt[k]{|a_k|} = a < 1$, dass ein $q < 1$ mit $a < q$ existiert, und wir können wieder das Vergleichskriterium aus Proposition 2.6.9 mit der geometrischen Reihe $\sum q^k$ anwenden. Für (2) folgt aus

$\limsup_{k\to\infty} \sqrt[k]{|a_k|} = a > 1$, dass für unendlich viele Folgeglieder a_k gilt $|a_k| > 1$, damit kann (a_k) keine Nullfolge sein. □

ÜBUNGSAUFGABE 2.6.14. *Für welche k konvergiert die Reihe $\sum_{n=2}^{\infty} \frac{n-2}{n^k}$?*
Hinweis: *Sie dürfen die Konvergenz der Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$ benutzen.*

ÜBUNGSAUFGABE 2.6.15. *Sei $\sum_{n=0}^{\infty} z_n$ eine komplexe Reihe. Zeigen Sie, dass die Reihe genau dann konvergiert, wenn die Reihen $\sum_{n=0}^{\infty} \mathrm{Re}(z)$ und $\sum_{n=0}^{\infty} \mathrm{Im}(z)$ konvergieren, und dass für die Summe gilt*

$$\sum_{n=0}^{\infty} z_n = \sum_{n=0}^{\infty} \mathrm{Re}(z) + \mathrm{i} \sum_{n=0}^{\infty} \mathrm{Im}(z)$$

ÜBUNGSAUFGABE 2.6.16. *Eine Folge komplexer Zahlen $(a_n)_n$ heißt* quadratsummierbar*, wenn $\sum_{n=1}^{\infty} \|a_n\|^2$ konvergiert. Zeigen Sie für quadratsummierbare Folgen $(a_n)_n$ und $(b_n)_n$ von komplexen Zahlen:*
(a) Die Reihe $\sum_{n=1}^{\infty} a_n b_n$ konvergiert absolut.
(b) Die Folge $(a_n + b_n)_n$ ist quadratsummierbar.

ÜBUNGSAUFGABE 2.6.17. *Bestimmen Sie in den folgenden Fällen, ob die Folgen $(a_n)_n$ bzw. die Reihen $\sum_{n=0}^{\infty} a_n$ konvergieren. Falls die Folge/Reihe konvergiert, bestimmen Sie Grenzwert bzw. Summe.*
(a) $a_n = \frac{4n^4+34n^2+7}{2n^4-12n^3+3n-5} + \frac{100n}{n^3-n^2}$.
(b) $b_n = \frac{5^n}{3^{2n}}$
(c) $c_n = \frac{1}{(n+11)(n+12)}$

ÜBUNGSAUFGABE 2.6.18. *Zeigen Sie, dass die folgenden Reihen konvergieren:*

$$\sum_{n=1}^{\infty} \frac{n!}{n^n}, \qquad \sum_{n=1}^{\infty} \left(\frac{n}{2n+1}\right)^n, \qquad \sum_{n=2}^{\infty} \frac{(-1)^n}{n-\sqrt{n}}, \qquad \sum_{n=1}^{\infty} n^{-n}.$$

DEFINITION 2.6.19. *Sei $\sum_{k=0}^{\infty} a_k$ eine Reihe. Für eine bijektive Abbildung*

$$\sigma \colon \mathbb{N} \to \mathbb{N}$$

heißt $\sum_{k=0}^{\infty} a_{\sigma(k)}$ eine Umordnung *der Reihe $\sum_{k=0}^{\infty} a_k$.*

BEMERKUNG 2.6.20. *Die umgeordnete Reihe hat dieselben Summanden wie die ursprüngliche Reihe, allerdings wird in anderer Reihenfolge aufsummiert. Für endliche Summen haben wir Kommutativität, d.h. der Wert einer endlichen Summe ist unabhängig von der Summationsreihenfolge. Bei Reihen haben wir unendlich viele Summanden und definieren den Wert als Grenzwert, darum ist nicht klar, dass die Umordnung wieder denselben Grenzwert hat.*

SATZ 2.6.21 (großer Unordnungssatz). *Jede Umordnung einer absolut konvergenten Reihe konvergiert absolut und hat denselben Grenzwert.*

BEWEIS. Sei $A = \sum_{n=0}^{\infty} a_n$ die Summe der Reihe und $\sigma \colon \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Wir wollen zeigen

$$\lim_{m\to\infty} \sum_{k=0}^{m} a_{\sigma(k)} = A.$$

Da die Reihe $\sum_{n=0}^{\infty} a_n$ absolut konvergiert, existiert für $\epsilon > 0$ ein N' so dass $\sum_{k=N'}^{\infty} |a_k| < \frac{\epsilon}{2}$. Dann haben wir mit Dreiecksungleichung

$$\left| A - \sum_{k=0}^{N'-1} a_k \right| = \left| \sum_{k=N'}^{\infty} a_k \right| \le \sum_{k=N'}^{\infty} |a_k| < \frac{\epsilon}{2}.$$

Da σ bijektiv ist, existiert ein $N \in \mathbb{N}$, so dass

$$\{\sigma(0), \sigma(1), \dots, \sigma(N)\} \supset \{0, 1, 2, \dots, N'-1\}.$$

Für alle $m \geq N$ haben wir die Menge

$$M = \{\sigma(0), \sigma(1), \dots, \sigma(m)\} \setminus \{0, 1, \dots, N'-1\} \subseteq \mathbb{N} \setminus \{0, 1, \dots, N'-1\}$$

und damit

$$\left|\sum_{k=0}^{m} a_{\sigma(k)} - \sum_{k=0}^{N'-1} a_k\right| = \left|\sum_{j \in M} a_j\right| \leq \sum_{j \in M} |a_j| \leq \sum_{k=N'}^{\infty} |a_k| \leq \frac{\epsilon}{2}.$$

Noch einmal die Dreiecksungleichung anwenden liefert

$$\left|\sum_{k=0}^{m} a_{\sigma(k)} - A\right| \leq \left|\sum_{k=0}^{m} a_{\sigma(k)} - \sum_{k=0}^{N'-1} a_k\right| + \left|\sum_{k=0}^{N'-1} a_k - A\right| < 2\frac{\epsilon}{2} = \epsilon.$$

Damit konvergiert die umgeordnete Reihe. Wenden wir das Argument auf $\sum_{k=0}^{\infty} |a_k|$ an, sehen wir auch die absolute Konvergenz der umgeordneten Reihe. □

Beispiel 2.6.22. *Die absolute Konvergenz ist notwendig. Zum Beispiel gibt es eine Umordnung der alternierenden harmonischen Reihe $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$, die divergiert: Für jedes n haben wir*

$$\frac{1}{2^n+1} + \frac{1}{2^n+3} + \dots + \frac{1}{2^{n+1}-1} > 2^{n-1}\frac{1}{2^{n+1}} = \frac{1}{4}.$$

Damit divergiert die Umordnung

$$\begin{aligned} 1 - \frac{1}{2} \quad &+ \quad \frac{1}{3} - \frac{1}{4} \\ &+ \quad \left(\frac{1}{5} + \frac{1}{7}\right) - \frac{1}{6} \\ &+ \quad \cdots \\ &+ \quad \left(\frac{1}{2^n+1} + \frac{1}{2^n+3} + \dots + \frac{1}{2^{n+1}-1}\right) - \frac{1}{2n+2} \\ &+ \quad \cdots \end{aligned}$$

□

Für nicht absolut konvergente Reihen zeigt die folgende Aufgabe, dass beim Umordnen alles Mögliche passieren kann.

Übungsaufgabe 2.6.23. *Sei $\sum_{k=0}^{\infty} a_k$ eine konvergente aber nicht absolut konvergente Reihe. Zeigen Sie die folgenden Aussagen:*

(1) *Für beliebig vorgegebenen Grenzwert $x \in \mathbb{R}$ eine Umordnung $\sum_{i=0}^{\infty} a_{\sigma(i)}$ existiert, deren Wert x ist.*
(2) *Es gibt Umordnungen der Reihe, die bestimmt gegen $+\infty$ bzw. $-\infty$ konvergieren.*
(3) *Es gibt Umordnungen der Reihe, die weder konvergieren noch bestimmt divergieren.*

Satz 2.6.24 (Cauchy-Produktsatz). *Sind $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ absolut konvergent, dann ist auch $\sum_{n=0}^{\infty} \sum_{i+j=n} a_i b_j$ absolut konvergent und es gilt*

$$\sum_{n=0}^{\infty} \sum_{i+j=n} a_i b_j = \left(\sum_{k=0}^{\infty} a_k\right) \cdot \left(\sum_{k=0}^{\infty} b_k\right).$$

Beweis. Wir beweisen erst die absolute Konvergenz einer Umordnung, und wenden dann den Unordnungssatz 2.6.21 an. Betrachten wir also die Reihe

$$a_0b_0 + a_0b_1 + a_1b_1 + a_1b_0 + \cdots + a_0b_N + \cdots + a_Nb_N + \cdots + a_Nb_0 + \cdots$$

Die Partialsumme mit Index $N^2 - 1$ für die Reihe der Absolutbeträge ist

$$\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} |a_k| \cdot |b_l| = \left(\sum_{k=0}^{N-1} |a_k|\right) \cdot \left(\sum_{l=0}^{N-1} |b_l|\right)$$

Nach Voraussetzung sind diese Partialsummen beschränkt (durch das Produkt oberer Schranken für $\sum_{k=0}^{N-1} |a_k|$ und $\sum_{l=0}^{N-1} |b_l|$). Das bedeutet, dass die angegebene Reihe absolut konvergiert. Aus der Beschreibung der Partialsumme mit Index N^2-1 sehen wir auch, dass diese Reihe gegen das Produkt der Grenzwerte der Reihen $\sum_{k=0}^{N-1} |a_k|$ und $\sum_{l=0}^{N-1} |b_l|$ konvergiert. Dafür wenden wir Proposition 2.2.1 auf die Teilfolge der Partialsummen mit Indizes $N^2 - 1$ an; dies ist die Folge der Produkte der Partialsummen der Reihen $\sum_{k=0}^{N-1} |a_k|$ und $\sum_{l=0}^{N-1} |b_l|$.

Nach dem Unordnungssatz 2.6.21 können wir die Reihe zu einer absolut konvergenten Reihe

$$a_0b_0 + a_0b_1 + a_1b_0 + \cdots + a_0b_m + a_1b_{m-1} + \cdots + a_mb_0 + \cdots$$

umordnen. Die Partialsummenfolge dieser Reihe hat die Partialsummenfolge der Reihe $\sum_{n=0}\sum_{i+j=n} a_ib_j$ als Teilfolge, die letztere Reihe ist also auch absolut konvergent. Nach dem Unordnungssatz ist der Grenzwert der selbe, wie für die oben betrachtete Reihe, damit ist auch die Aussage zum Wert dieser Reihe bewiesen. □

Beispiel 2.6.25.
- *Eine Anwendung des Cauchy-Produktsatzes ist die Funktionalgleichung für die Exponentialfunktion, s. Abschnitt 5.1.*
- *Die geometrische Reihe $\sum_{k=0}^{\infty} z^k$ konvergiert für $\|z\| < 1$ absolut gegen $\frac{1}{1-z}$. Aus dem Cauchy-Produktsatz folgt dann*

$$\frac{1}{(1-z)^2} = \sum_{n=0}^{\infty}(n+1)z^n.$$

□

2.7. Anwendungen

2.7.1. Dezimalbruchdarstellung, b-adische Brüche. Für eine natürliche Zahl $b \geq 2$ (wie *Basis*) definieren wir einen *b-adische Bruch* als unendliche Reihe

$$\pm z_{-n}z_{-n+1}\dots z_0, z_1z_2z_3\cdots := \pm\sum_{k=-n}^{\infty}\frac{z_k}{b^k}.$$

Die Ziffern sind dabei ganze Zahlen mit $0 \leq z_i \leq b-1$. Im Spezialfall $b = 10$ erhalten wir die üblichen Dezimalbrüche, für $b = 2$ *dyadische Brüche.* Die Ziffern $z_{-n}, \dots, z_0$ werden üblicherweise vor dem Komma, und die Ziffern $z_1, \dots, z_n, \dots$ nach dem Komma geschrieben.

Betrachten wir nur b-adische Brüche der Form $\sum_{k=1}^{\infty}\frac{z_k}{b^k}$. Dann bilden die Partialsummen

$$0, z_1z_2\dots z_n := \sum_{k=1}^{n}\frac{z_k}{b^k}$$

eine monoton wachsende Folge nicht-negativer Zahlen, die durch $\sum_{k=1}^{\infty}\frac{b-1}{b^k} = 1$ beschränkt ist. (Hier versteckt sich das berühmte $0,999\cdots = 1$ oder das weniger berümte dyadische $0,1111\cdots = 1$ in Form einer geometrischen Reihe.) Die unendliche Reihe konvergiert also und stellt eine eindeutige Zahl im Intervall $[0,1]$ dar.

Umgekehrt können wir jede reelle Zahl eindeutig als b-adischen Bruch entwickeln, d.h. für jede reelle Zahl x gibt es eindeutige natürliche Zahlen $z_{-n}, z_{-n+1}, \dots, z_0, z_1, \dots$ mit $0 \leq z_i < b$, so dass

$$\pm z_{-n} \dots z_0, z_1 z_2 \dots z_k \leq x < \pm z_{-n} \dots z_0, z_1 z_2 \dots z_k + \frac{1}{b^k}.$$

Diese können wir rekursiv bestimmen. Nehmen wir an, dass $x \geq 0$ (andernfalls haben wir nur ein Vorzeichen zu ändern). Für den Rekursionsanfang ist n die natürliche Zahl, so dass $b^n \leq x < b^{n+1}$ ist, und $z_{-n} = \lfloor \frac{x}{b^n} \rfloor$. Haben wir die Ziffern $z_{-n}, \dots, z_0, z_1, \dots, z_{k-1}$ bestimmt, ist

$$z_k = \lfloor b^k (x - z_{-n} \dots z_0, z_1 z_2 \dots z_{k-1}) \rfloor$$

Mit vollständiger Induktion sehen wir, dass $0 \leq z_k < b$ und $z_{-n} \dots z_0, z_1 z_2 \dots z_k \leq x < z_{-n} \dots z_0, z_1 z_2 \dots z_k + \frac{1}{b^k}$ gelten. Damit konvergiert der b-adische Bruch

$$\sum_{k=-n}^{\infty} \frac{z_k}{b^k}$$

gegen x.

Bemerkung 2.7.1. *Alternativ kann die reelle Zahl x auch als Supremum der Menge der Approximationen $\{z_{-n} \dots z_0, z_1 z_2 \dots z_k \mid k \in \mathbb{Z}, k \geq -n\}$ oder durch die Intervallschachtelung*

$$[z_{-n} \dots z_0, z_1 z_2 \dots z_k ; z_{-n} \dots z_0, z_1 z_2 \dots z_k + \frac{1}{b^k}], \quad k \geq -n$$

definiert werden.

Allgemein ist die Beschreibung reeller Zahlen durch Dezimalbrüche (oder allgemeiner b-adische Brüche) problematisch, da für den Dezimalbruch unendlich viele Nachkommastellen angegeben werden müssen. Die einzige Möglichkeit dafür ist, eine Maschine/ein Programm anzugeben, dass nacheinander die Nachkommastellen auswirft (und dafür natürlich unendlich lange laufen muss). Die reellen Zahlen, die auf diese Art beschrieben werden können, heißen berechenbare reelle Zahlen*. Da die Programme, die die Zahlen beschreiben, endlich sind, ist die Menge der Programme, und damit die Menge der berechenbaren reellen Zahlen, abzählbar. Es kann also nur eine verschwindend geringe Menge von reellen Zahlen auf diese Art beschrieben werden.*

2.7.2. Zeta-Funktion.

Proposition 2.7.2. *Für $s \in \mathbb{Q}$ ist die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^s}$ absolut konvergent, falls $s > 1$ ist und divergent, falls $s \leq 1$ ist.*

Beweis. Sei $s > 1$. Da $\frac{1}{n^s} > 0$, können wir die Partialsummen S_n durch Partialsummen S_{2^m-1} und eine geometrische Summe abschätzen:

$$\begin{aligned} S_n \leq S_{2^m-1} &= 1 + \left(\frac{1}{2^s} + \frac{1}{3^s}\right) + \dots + \left(\frac{1}{2^{(m-1)s}} + \dots + \frac{1}{(2^m-1)^s}\right) \\ &\leq 1 + 2 \cdot \frac{1}{2^s} + \dots + 2^{m-1} \cdot \frac{1}{2^{(m-1)s}} \\ &= \frac{1 - 2^{(1-s)m}}{1 - 2^{1-s}} < \frac{1}{1 - 2^{1-s}}. \end{aligned}$$

Damit ist die Folge der Partialsummen $\left(\sum_{k=1}^{n} \frac{1}{n^s}\right)$ beschränkt und damit absolut konvergent, nach Proposition 2.6.7.

Sei $s \leq 1$. Dann ist

$$S_n = 1 + \frac{1}{2^s} + \cdots + \frac{1}{n^s} \geq \sum_{k=1}^{n} \frac{1}{k}.$$

Nach Proposition 2.6.9 divergiert die Reihe, da die harmonische Reihe divergiert. □

Die Definition lässt sich auf komplexe Zahlen $s \in \mathbb{C}$ erweitern, wenn wir die allgemeine Potenz diskutieren, s. Abschnitt 5.1. Die entstehende Funktion

$$\zeta(s) = \sum_{n=0}^{\infty} \frac{1}{n^s}, \quad s \in \mathbb{C}$$

heißt *Riemannsche Zeta-Funktion* und spielt in der Zahlentheorie eine herausgehobene Rolle.

PROPOSITION 2.7.3. *Für n bezeichnen wir mit $\mathcal{P}_n$ die Menge der ersten n Primzahlen und mit $\mathbb{N}(\mathcal{P}_n)$ die Menge der natürlichen Zahlen, in deren Primfaktorisierung nur die Primzahlen aus $\mathcal{P}_n$ vorkommen. Dann gilt für alle n und alle rationalen $s > 1$:*

$$\prod_{p \in \mathcal{P}_n} \frac{1}{1 - p^{-s}} = \sum_{N \in \mathbb{N}(\mathcal{P}_n)} \frac{1}{N^s}.$$

Für $n \to \infty$ erhalten wir die Euler-Produktdarstellung der Zeta-Funktion

$$\zeta(s) = \prod_{p \text{ prim}} \frac{1}{1 - p^{-s}}$$

BEWEIS. Der Beweis erfolgt durch Induktion über n. Für den Induktionsanfang haben wir aus der geometrischen Reihe

$$\frac{1}{1 - 2^{-s}} = \sum_{k=0}^{\infty} \frac{1}{2^{ks}} = \sum_{N \in \mathbb{N}(\{2\})} \frac{1}{N^s}.$$

Im Induktionsschritt haben wir

$$\begin{aligned}
\prod_{p \in \mathcal{P}_{n+1}} \frac{1}{1 - p^{-s}} &= \left(\prod_{p \in \mathcal{P}_n} \frac{1}{1 - p^{-s}} \right) \frac{1}{1 - p_{n+1}^{-s}} \\
&= \left(\sum_{N \in \mathbb{N}(\mathcal{P}_n)} \frac{1}{N^s} \right) \left(\sum_{k=0}^{\infty} \frac{1}{p_{n+1}^{ks}} \right) \\
&= \sum_{N \in \mathbb{N}(\mathcal{P}_{n+1})} \frac{1}{N^s}
\end{aligned}$$

durch Anwendung der Induktionsbehauptung und der geometrischen Reihe. Multiplikation und Umordnung im letzten Schritt sind zulässig, da die Reihen absolut konvergieren. □

BEMERKUNG 2.7.4.
- *Im Beweis wurden Existenz und Eindeutigkeit der Primfaktorzerlegung für natürliche Zahlen benutzt.*
- *Die Konvergenz des Euler-Produkts $\prod_{p \text{ prim}} \frac{1}{1-p^{-s}}$ kann auch unabhängig von der Identifikation mit der Zeta-Funktion bewiesen werden.*

ÜBUNGSAUFGABE 2.7.5. *(a) Benutzen Sie die Zerlegung $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$, um den Grenzwert der Reihe $\sum_{n=2}^{\infty} \frac{1}{n(n+1)}$ zu bestimmen.*

(b) Benutzen Sie die Reihe aus Teil (a), um die folgende Abschätzung zu zeigen:

$$\frac{3}{2} \leq \sum_{n=1}^{\infty} \frac{1}{n^2} \leq 2$$

KAPITEL 3

Funktionenfolgen und Funktionenreihen

3.1. Gleichmäßige Konvergenz

Wir können auch Folgen von reell-wertigen oder komplex-wertigen Funktionen betrachten. Bezeichnen wir für einen festen Definitionsbereich D mit $\mathrm{Fun}(D, \mathbb{R})$ die reell-wertigen Funktionen, dann ist eine Folge von Funktionen eine Abbildung $\mathbb{N} \to \mathrm{Fun}(D, \mathbb{R})$, die wir als $(f_n)_{n\in\mathbb{N}}$ schreiben.

Definition 3.1.1. *Sei D eine Menge.*

(1) *Eine Funktionenfolge $(f_n : D \to \mathbb{R})_{n\in\mathbb{N}}$* konvergiert punktweise *gegen eine Funktion $f : D \to \mathbb{R}$, wenn für alle $x \in D$ gilt $f(x) = \lim_{n\to\infty} f_n(x)$. Mit Quantoren:*
$$\forall x \in D\, \forall \epsilon > 0\, \exists N \in \mathbb{N}\, \forall n \geq N : |f_n(x) - f(x)| < \epsilon.$$
(2) *Eine Funktionenfolge $(f_n : D \to \mathbb{R})_{n\in\mathbb{N}}$* konvergiert gleichmäßig *gegen eine Funktion $f : D \to \mathbb{R}$, wenn*
$$\forall \epsilon > 0\, \exists N \in \mathbb{N}\, \forall n \geq N\, \forall x \in D : |f_n(x) - f(x)| < \epsilon.$$

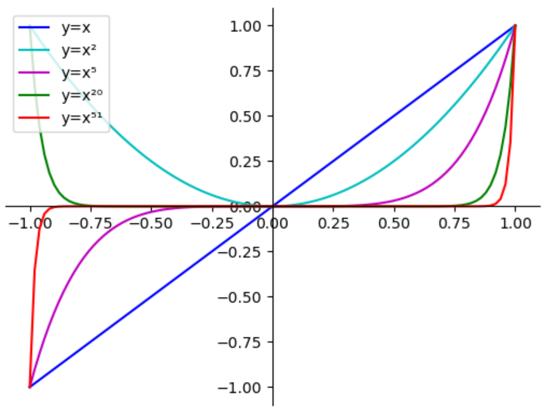


Abbildung 7. Funktionenfolge $(x^n)_{n\in\mathbb{N}}$ auf $[-1, 1]$

Beispiel 3.1.2. *Wir betrachten die Funktionenfolge $(x^n)_{n\in\mathbb{N}}$. Auf $(-1, 1]$ konvergiert diese Folge punktweise gegen die Grenzfunktion*
$$f(x) = \begin{cases} 0 & x \in (-1, 1) \\ 1 & x = 1 \end{cases}$$
Dies ist auch der maximale Bereich, auf dem wir punktweise Konvergenz haben, das folgt aus der Konvergenzbetrachtung der geometrischen Folge s. Beispiel 2.1.5.

Die Konvergenz ist auf $[0, 1]$ nicht gleichmäßig. Dafür müssen wir zeigen, dass ein ϵ existiert, so dass für beliebig große n gilt $x^n > \epsilon$. Für irgendein $0 < \epsilon < 1$ und *beliebig großes* n *können wir immer ein* x *mit* $\sqrt[n]{\epsilon} \leq x < 1$ *finden, also haben wir keine gleichmäßige Konvergenz.*

Betrachten wir $0 < a < 1$, *dann ist die Konvergenz auf* $[0, a]$ *gleichmäßig. Da* x^n *für* $0 \leq x < 1$ *eine Nullfolge ist, existiert für* $\epsilon > 0$ *ein* N, *so dass* $a^n < \epsilon$ *für alle* $n > N$. *Aus der Monotonie der Potenzen folgt dann* $x^n \leq a^n < \epsilon$ *für alle* $x \in [0, a]$. *Damit haben wir gleichmäßige Konvergenz.* □

BEMERKUNG 3.1.3.
- *Analog definieren die punktweise und gleichmäßige Konvergenz für komplex-wertige Funktionen.*
- *Punktweise und gleichmäßige Konvergenz unterscheiden sich in der Reihenfolge der Quantoren für* $x \in D$ *und* N. *Bei punktweiser Konvergenz kann das* N, *ab dem die Funktionen auf dem Definitionsbereich näher als* ϵ *an der Grenzfunktion liegen, von* $x \in D$ *abhängig sein. Bei gleichmäßiger Konvergenz existiert ein* N, *ab dem die Funktionen* f_n auf dem gesamten Definitionsbereich ϵ*-nah an der Grenzfunktion sind.*
- *Wir können für (reell- oder komplex-wertige) Funktionen auf dem Definitionsbereich* D *auch die* Supremums-Norm *definieren:*

$$\|f\|_\infty := \sup\{|f(x)| \mid x \in D\}$$

Damit kann die gleichmäßige Konvergenz umformuliert werden:

$$\forall \epsilon > 0 \, \exists N \in \mathbb{N} \, \forall n \geq N \colon \|f_n - f\|_\infty < \epsilon.$$

- *Wir werden später sehen, dass gleichmäßige Konvergenz sehr nützlich ist, da sich Eigenschaften wie Stetigkeit, Differenzierbarkeit und Integrierbarkeit auf Grenzwerte gleichmäßig konvergenter Funktionenfolgen übertragen (bzw. auch Grenzprozesse wie ableiten und integrieren mit Grenzwerten gleichmäßig konvergenter Funktionenfolgen vertauschen).*

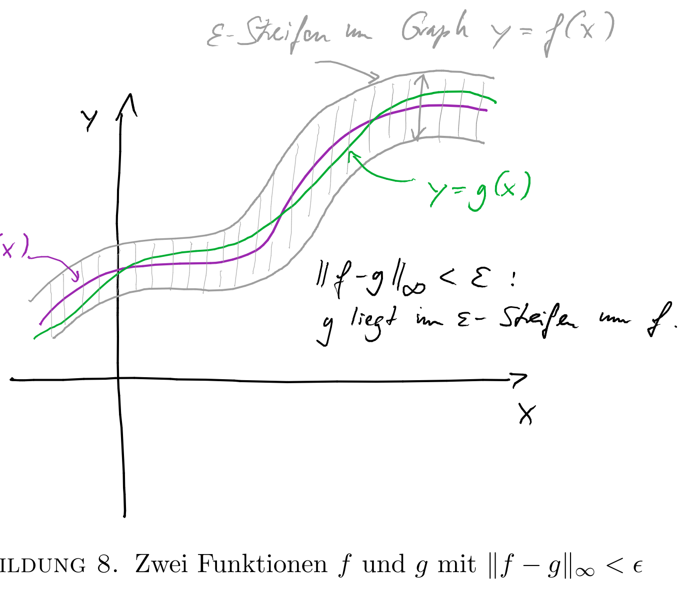


ABBILDUNG 8. Zwei Funktionen f und g mit $\|f - g\|_\infty < \epsilon$

ÜBUNGSAUFGABE 3.1.4. *Zeigen Sie, dass die Supremums-Norm auf der Menge der Funktionen* $f \colon D \to \mathbb{R}$ *die folgenden Eigenschaften hat:*

positive Definitheit: $\|f\| \geq 0$ *für alle Funktionen* $f \colon D \to \mathbb{R}$ *und* $\|f\| = 0$ *genau dann, wenn* $f \equiv 0$.

Skalarmultiplikation: $\|\lambda \cdot f\| = \|\lambda\| \cdot \|f\|$ *für alle reellen Zahlen* $\lambda \in \mathbb{R}$ *und alle Funktionen* $f\colon D \to \mathbb{R}$ *mit* $\|f\|_\infty < \infty$.

Dreiecks-Ungleichung: $\|f+g\| \leq \|f\| + \|g\|$ *für alle* $f, g\colon D \to \mathbb{R}$.

DEFINITION 3.1.5. *Gegeben eine Folge* $(f_n\colon D \to \mathbb{R})_{n\in\mathbb{N}}$ *von Funktionen auf dem Definitionsbereich* D *können wir auch die unendliche Reihe*

$$\sum_{k=0}^{\infty} f_k$$

betrachten. Die Reihe konvergiert (absolut bzw. gleichmäßig), wenn die Funktionenfolge $(\sum_{k=0}^{n} f_k)_{n\in\mathbb{N}}$ *der Partialsummen (absolut bzw. gleichmäßig) konvergiert.*

BEISPIEL 3.1.6. *Wir betrachten* $f_n\colon (-1,1) \to \mathbb{R}\colon x \mapsto \sum_{k=0}^{n} x^k$. *Für jeden Punkt* $x \in (-1,1)$ *haben wir eine geometrische Reihe, die konvergiert. Damit konvergiert die Folge* $(f_n)_{n\in\mathbb{N}}$ punktweise *gegen die Grenzfunktion* $f(x) = \frac{1}{1-x}$.

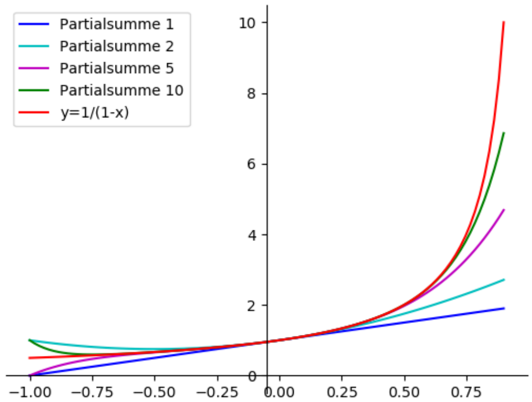


ABBILDUNG 9. Geometrische Reihe und Grenzfunktion auf $(-1, 0.85)$

Die Konvergenz ist aber nicht gleichmäßig, wir zeigen die Negation der gleichmäßigen Konvergenz:

$$\exists \epsilon > 0\, \forall N \in \mathbb{N}\, \exists n > N \exists x \in (-1,1)\colon |f_n(x) - f(x)| \geq \epsilon.$$

Die Ungleichung ist in unserem Fall

$$|f_n(x) - f(x)| = \left|\sum_{k=0}^{n} x^k - \frac{1}{1-x}\right| = \left|\frac{1-x^{n+1}}{1-x} - \frac{1}{1-x}\right| = \frac{|x|^{n+1}}{|1-x|} \geq \epsilon.$$

Wir setzen $\epsilon = 1$. *Dann müssen wir zeigen, dass für jedes* N *ein* $n > N$ *und ein* $x \in (-1,1)$ *existiert, so dass* $|x|^{n+1} \geq |1-x|$. *Für* $0 < \delta < 1$ *und* $x = 1-\delta$ *haben wir* $|x| = x$ *und* $|1-x| = 1-x$, *und die Ungleichung* $x^{n+1} \geq 1-x$ *ist äquivalent zu* $(1-\delta)^{n+1} + (1-\delta) - 1 = (1-\delta)^{n+1} - \delta \geq 0$. *Mit der Bernoulli-Ungleichung schätzen wir ab:*

$$(1-\delta)^{n+1} - \delta \geq 1 - (n+1)\delta - \delta = 1 - (n+2)\delta.$$

Das bedeutet, dass für $\delta < \frac{1}{n+2}$ *und* $x = 1-\delta$ *die Ungleichung* $|x|^{n+1} \geq |1-x|$ *erfüllt ist, und wir haben gezeigt, dass die Konvergenz nicht gleichmäßig ist.* □

Wir diskutieren noch kurz einige Kriterien für gleichmäßige Konvergenz.

PROPOSITION 3.1.7 (Cauchy-Kriterium). *Eine Folge $f_n\colon D \to \mathbb{R}$ von Funktionen auf dem Definitionsbereich D konvergiert genau dann gleichmäßig, wenn es zu jedem $\epsilon > 0$ eine Zahl N gibt, so dass $\|f_n - f_m\|_\infty < \epsilon$ für alle $n, m \geq N$ gilt.*

BEWEIS. Wenn (f_n) gleichmäßig gegen f konvergiert, gibt es nach Definition für jedes $\epsilon > 0$ ein N, so dass $\|f_n - f\|_\infty < \frac{\epsilon}{2}$ für $n > N$. Mit der Dreiecksungleichung haben wir dann für alle $n, m > N$

$$\|f_n - f_m\|_\infty \leq \|f_n - f\|_\infty + \|f - f_m\|_\infty < \epsilon.$$

Wenn umgekehrt die Bedingung erfüllt ist, folgt aus $|f_n(x) - f_m(x)| < \epsilon$ für alle $x \in D$ und alle $n, m > N$, dass $(f_n(x))_{n\in\mathbb{N}}$ eine Cauchy-Folge ist. Damit haben wir punktweise Konvergenz gegen eine Grenzfunktion $f(x) := \lim_{n\to\infty} f_n(x)$. Aus der Bedingung folgt (mit $m \to \infty$) dann auch $|f_n(x) - f(x)| < \epsilon$ für alle $x \in D$ und $n > N$. □

SATZ 3.1.8 (Weierstraß-Majorantenkriterium). *Sei $D \subset \mathbb{R}$. Gilt für jede Funktion $f_n\colon D \to \mathbb{R}$ der Folge $(f_n)_{n\in\mathbb{N}}$ eine Abschätzung $|f_n(x)| \leq M_n$ für alle $x \in D$ (alternativ: $\|f_n\| \leq M_n$), so dass $\sum_{n=0}^\infty M_n$ konvergiert, dann ist*

$$\sum_{n=0}^\infty f_n$$

auf D gleichmäßig und absolut konvergent.

BEWEIS. Wir zeigen erst die punktweise Konvergenz gegen eine Grenzfunktion, und zeigen dann, dass die Konvergenz gleichmäßig ist.

Für die punktweise Konvergenz sei $x \in D$. Aus $|f_n(x)| \leq M_n$ und dem Majorantenkriterium 2.6.9 haben wir die absolute Konvergenz von $\sum_{n=0}^\infty f_n(x)$. Damit erhalten wir eine Funktion

$$F(x) := \sum_{n=0}^\infty f_n(x).$$

Wir betrachten die Folge der Partialsummenfunktionen $S_n = \sum_{k=0}^n f_k$ und wollen zeigen, dass diese gleichmäßig gegen F konvergiert. Für $\epsilon > 0$ folgt aus der Konvergenz von $\sum_{k=0}^\infty M_n$, dass ein N existiert, so dass

$$|S_n(x) - F(x)| = \left|\sum_{k=n+1}^\infty f_k(x)\right| \leq \sum_{k=n+1}^\infty |f_k(x)| \leq \sum_{k=n+1}^\infty M_k < \epsilon$$

für alle $n > N$ und alle $x \in D$. Das ist die gleichmäßige Konvergenz. □

BEISPIEL 3.1.9. *Die Reihe $\sum_{n=1}^\infty \frac{\cos(nx)}{n^2}$ konvergiert auf $\mathbb{R}$ gleichmäßig, da $\|f_n\|_\infty = \frac{1}{n^2}$ gilt und $\sum_{n=1}^\infty \frac{1}{n^2}$ konvergiert.* □

BEMERKUNG 3.1.10. *Funktionenreihen $\sum_{k=0}^\infty f_k$ von Funktionen $f_n\colon D \to \mathbb{R}$, so dass alle f_n beschränkt sind und $\sum_{n=0}^\infty \|f_n\|_\infty$ konvergiert, heißen* normal konvergent. *Nach dem Weierstraß-Kriterium aus Satz 3.1.8 sind normal konvergente Funktionenreihen gleichmäßig konvergent.*

Es gibt auch noch andere Kriterien für die gleichmäßige Konvergenz von Funktionenreihen (Dirichlet oder Abel), die wir hier aber nicht diskutieren werden.

ÜBUNGSAUFGABE 3.1.11. *Zeigen Sie, dass die Funktionenreihe $\sum_{n=1}^\infty \frac{1}{n(x-n)}$ für jedes abgeschlossene Intervall $[a, b] \subseteq \mathbb{R} \setminus \mathbb{N}$ gleichmäßig konvergiert.*

Hinweis: *Sie dürfen benutzen, dass $\sum_{n=1}^\infty \frac{1}{n^2}$ konvergiert.*

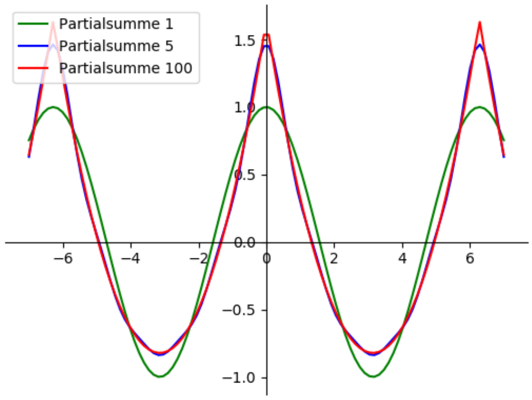


ABBILDUNG 10. Partialsummen der Funktionenreihe $\sum_{n=1}^{\infty} \frac{\cos(nx)}{n^2}$

ÜBUNGSAUFGABE 3.1.12. *Sei $(f_n)_n$ eine gleichmäßig konvergente Folge von Funktionen auf D. Wir nehmen an, dass ein $a > 0$ existiert, so dass $|f_n(x)| \geq a$ für alle $n \in \mathbb{N}$ und $x \in D$ gilt. Zeigen Sie, dass die Folge $(\frac{1}{f_n})_n$ gleichmäßig gegen $\frac{1}{f}$ konvergiert, wobei $f = \lim_{n\to\infty} f_n$ ist.*

3.2. Potenzreihen

DEFINITION 3.2.1. *Funktionenreihen der Form $\sum_{k=0}^{\infty} a_k x^k$ mit $a_k \in \mathbb{R}$ bzw. $a_k \in \mathbb{C}$ heißen* Potenzreihen *(mit reellen bzw. komplexen Koeffizienten).*

BEMERKUNG 3.2.2. *Damit können wir Funktionen als Grenzwert einer Folge von Polynomen definieren. Wir werden in den Abschnitten zu Stetigkeit, Differenzierbarkeit und Integrierbarkeit untersuchen, wie sich die Eigenschaften von Potenzreihen auf Eigenschaften von Polynomen zurückführen lassen. Im Abschnitt zur Taylorentwicklung werden wir Potenzreihenentwicklungen für eine große Klasse von Funktionen diskutieren.*

Die selbe Idee funktioniert für Reihen von trigonometrischen Funktionen, s. Fouriertheorie in Analysis 2 oder 3.

DEFINITION 3.2.3. *Sei $\sum_{k=0}^{\infty} a_k x^k$ eine Potenzreihe. Wir betrachten die Menge $M = \{x \in \mathbb{R} \mid \sum_{k=0}^{\infty} a_k x^k$ ist konvergent$\}$. Der* Konvergenzradius *der Potenzreihe ist definiert als*

$$R = \begin{cases} \sup\{|x| \mid x \in M\} & \text{falls } M \text{ beschränkt} \\ \infty & \text{sonst} \end{cases}$$

BEMERKUNG 3.2.4. *Analog definieren wir den Konvergenzradius einer komplexen Potenzreihe.*

ÜBUNGSAUFGABE 3.2.5. *Zeigen Sie mit dem Majorantenkriterium die folgende alternative Beschreibung*

$$R = \sup\{r \geq 0 \mid (|a_n| r^n)_{n\in\mathbb{N}} \text{ ist beschränkt}\}$$

SATZ 3.2.6. *Sei $\sum_{k=0}^{\infty} a_k z^k$ eine Potenzreihe mit Konvergenzradius R. Dann gelten die folgenden Aussagen:*

(1) $R = 0$ gilt genau dann, wenn die Reihe nur für $z = 0$ (gegen a_0) konvergiert.

(2) Für $R > 0$ und $\rho \in (0, R)$ konvergiert die Reihe absolut und gleichmäßig auf $\|z\| \leq \rho$.
(3) Für z mit $\|z\| > R$ ist die Reihe divergent.

Beweis. (1) folgt direkt aus der Definition des Konvergenzradius und der Tatsache, dass $\|z\| = 0$ genau dann, wenn $z = 0$ ist.

(2) Aus der Annahme folgt, dass es z_0 mit $\rho < \|z_0\| \leq R$ gibt, so dass die Reihe konvergiert. Insbesondere ist $(\|a_k z_0^k\|)_{k \in \mathbb{N}}$ eine Nullfolge, und damit beschränkt durch eine Zahl c. Für $\|z\| \leq \rho$ und $k \in \mathbb{N}$ haben wir

$$\|a_k z^k\| = \|a_k z_0^k\| \cdot \left\| \frac{z}{z_0} \right\|^k \leq c \cdot \left\| \frac{z}{z_0} \right\|^k .$$

Mit $\left\| \frac{z}{z_0} \right\| < 1$ folgt die absolute und gleichmäßige Konvergenz aus Satz 3.1.8

(3) Sei $\|z\| > R$. Wenn $\sum_{k=0}^{\infty} a_k z^k$ konvergiert, dann ist nach (2) auch $\sum_{k=0}^{\infty} a_k r^k$ für jedes $R < r < \|z\|$ konvergent. Das ist ein Widerspruch zur Definition von R als Supremum. □

Bemerkung 3.2.7. *Wir machen hier keine Aussage über den Rand des Konvergenzgebiets $|z| = R$.*

Proposition 3.2.8. *Sei $\sum_{k=0}^{\infty} a_k z^k$ eine Potenzreihe.*

(1) Wir nehmen an, dass $a_k \neq 0$ für fast alle k ist und dass

$$a = \lim_{n \to \infty} \left\| \frac{a_{n+1}}{a_n} \right\|$$

existiert. Dann ist der Konvergenzradius der Reihe

$$R = \begin{cases} \frac{1}{a} & a \neq 0 \\ \infty & a = 0 \end{cases}$$

(2) Wenn $L = \limsup_n \sqrt[n]{\|a_n\|}$ existiert, dann ist der Konvergenzradius der Reihe

$$R = \begin{cases} \frac{1}{L} & L \neq 0 \\ \infty & L = 0 \end{cases}$$

Beweis. (2) Für $z \neq 0$ gilt

$$\limsup_n \sqrt[n]{\|a_n z^n\|} = \|z\| \cdot \limsup_n \sqrt[n]{\|a_n\|} \quad \begin{cases} < 1 & \text{wenn } \|z\| < \frac{1}{L} \\ > 1 & \text{wenn } \|z\| > \frac{1}{L} \end{cases}$$

Die Aussage folgt dann aus dem Wurzelkriterium 2.6.13.

(1) folgt analog aus dem Quotientenkriterium 2.6.11. □

Übungsaufgabe 3.2.9. *Bestimmen Sie den Konvergenzradius R der folgenden Potenzreihen*

$$\sum_{n=0}^{\infty} \frac{1}{n+1} \binom{2n}{n} x^n, \qquad \sum_{n=1}^{\infty} \frac{x^n}{(4n)^7}, \qquad \sum_{n=0}^{\infty} \frac{x^{3n}}{2^n}, \qquad \sum_{n=0}^{\infty} \frac{x^{2n+1}}{2n+1}, \qquad \sum_{n=0}^{\infty} (-1)^n n! x^{2n}$$

Eine direkte Folgerung aus dem Cauchy-Produktsatz 2.6.24 ist die folgende Aussage:

Proposition 3.2.10. *Seien $f(z) = \sum_{k=0}^{\infty} a_n z^n$ und $g(z) = \sum_{l=0}^{\infty} b_l z^l$ zwei Potenzreihen, die am Punkt z absolut konvergent sind. Dann gilt*

$$f(z) \cdot g(z) = \sum_{n=0}^{\infty} \left(\sum_{p+q=n} a_p b_q \right) z^n$$

LEMMA 3.2.11. *Sei $\sum_{k=0}^{\infty} a_k z^k$ eine Potenzreihe mit Konvergenzradius $R > 0$. Dann gibt es zu jedem positiven $r < R$ eine Konstante c, so dass für alle $\|z\| \leq r$ gilt:*

$$\left\| \sum_{k=n}^{\infty} a_k z^k \right\| \leq c\|z\|^n$$

BEWEIS. Für die Konstante wählen wir $c = \sum_{k=0}^{\infty} \|a_{n+k}\| r^k$. Dann gilt

$$\left\| \sum_{k=n}^{\infty} a_k z^k \right\| \leq \sum_{k=n}^{\infty} \|a_k\| \cdot \|z\|^k \leq c\|z\|^n.$$

□

PROPOSITION 3.2.12. *Sei $\sum_{k=0}^{\infty} a_k z^k \neq 0$ eine Potenzreihe mit Konvergenzradius $R > 0$. Dann gibt es einen Kreis um 0, der höchstens endlich viele Nullstellen von f enthält.*

BEWEIS. Sei N der kleinste Index mit $a_N \neq 0$. Zu einem beliebigen $r < R$ wählen wir wie in Lemma 3.2.11 ein c, so dass für alle $\|z\| \leq r$ gilt

$$\|f(z) - a_N z^N\| \leq c\|z\|^{N+1}.$$

Nehmen wir an, dass es unendlich viele Nullstellen gibt. Dann enthält jeder der Kreise

$$\{z \in \mathbb{C} \mid \|z\| \leq \frac{r}{k}\}$$

eine Nullstelle $z_k \neq 0$. Aus der obigen Abschätzung ergibt sich dann mit $f(z_k) = 0$ auch $\|a_n\| \leq c\|z_k\|$. Wegen $\lim_{k\to\infty} z_k = 0$ folgt dann $a_N = 0$, was ein Widerspruch ist. □

KOROLLAR 3.2.13 (Identitätssatz für Potenzreihen). *Seien $f(z) = \sum_{k=0}^{\infty} a_k z^k$ und $g(z) = \sum_{l=0}^{\infty} b_l z^l$ zwei Potenzreihen mit Konvergenzradius $\neq 0$. Wenn eine Nullfolge (z_k) von Zahlen $z_k \neq 0$ existiert, so dass $f(z_k) = g(z_k)$ für alle k gilt, dann gilt auch schon $a_k = b_k$ für alle k.*

BEMERKUNG 3.2.14. *erzeugende Funktionen*

ÜBUNGSAUFGABE 3.2.15. *rekursive Formel für Koeffizienten der inversen Potenzreihe. Konvergenzradius bis zur ersten Nullstelle. Beispiele Invertieren von $1 - x$, $1 - x - x^2$.*

ÜBUNGSAUFGABE 3.2.16. *Die Catalan-Zahlen C_n sind eine Folge natürlicher Zahlen aus der Kombinatorik. Die Catalan-Zahl C_n zählt dabei die Anzahl der möglichen vollständigen Klammerungen eines Produkts mit $n + 1$ Faktoren. (Alternativ ist das gleich der Anzahl der Binärbäume mit $n + 1$ Blättern, evtl. aus Informatik-Veranstaltungen bekannt. Es gibt weitere äquivalente Zählprobleme, bei denen die Catalan-Zahlen auftreten.) Zum Beispiel ist $C_3 = 5$:*

$$((ab)c)d, \quad (a(bc))d, \quad a((bc)d), \quad a(b(cd)), \quad (ab)(cd).$$

(a) Zeigen Sie, dass die Catalan-Zahlen die folgende Rekursionsgleichung erfüllen:

$$C_0 = 1, \quad C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}$$

(b) Zeigen Sie mit der Cauchy-Produktformel, dass die Potenzreihe $c(x) = \sum_{n=0}^{\infty} C_n x^n$ die folgende Gleichung erfüllt:

$$c(x) = 1 + xc(x)^2.$$

(c) *Benutzen Sie die quadratische Lösungsformel, um eine explizite Beschreibung der Funktion $c(x)$ zu erhalten. (Hier reicht es, die beiden Lösungen der quadratischen Gleichung anzugeben. Um eine der beiden Lösungen auszuschließen, muss $c(x)$ für $x \to 0$ untersucht werden, dafür benötigen wir aber Stetigkeitsargumente.)*

Bemerkung: Diese Überlegungen werden benutzt, um eine explizite (nicht-rekursive) Formel für die Catalan-Zahlen zu erhalten. Um dies zu erreichen, benötigen wir allerdings die Potenzreihendarstellung der Wurzelfunktion.

KAPITEL 4

Stetigkeit

4.1. Stetigkeit: Definitionen und Beispiele

Wir benutzen häufig Funktionen zur mathematischen Modellierung der Welt:

- die Periode/Schwingungsdauer eines Fadenpendels (bei kleinen Auslenkungen) als Funktion von der Fadenlänge,
- die Entfernung, in der ein Objekt nach einer parabolischen Flugbahn auf dem Boden auftrifft, als Funktion in Abhängigkeit von der Anfangsgeschwindigkeit oder vom Abwurfwinkel,
- der Weg, den ein Wagen nach einer Beschleunigung auf einer schiefen Ebene zurücklegt, bis er durch Rollreibung zum Stillstand kommt.

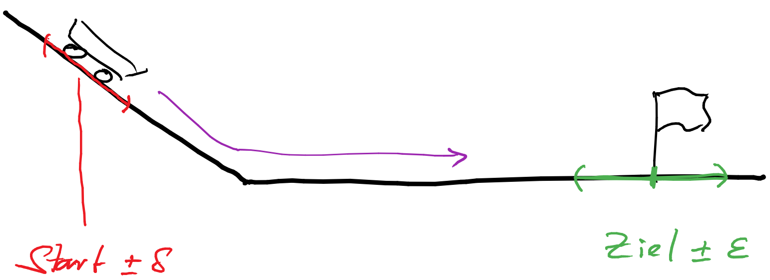


ABBILDUNG 11. Stetigkeit: Für ein gewünschtes Ergebnis $\pm\epsilon$ sind Startbedingungen $\pm\delta$ erforderlich.

In all diesen Fällen sehen wir, dass kleine Änderungen der Anfangsbedingungen auch nur beschränkte Änderungen im Ergebnis zur Folge haben. Tatsächlich ist nur dadurch eine sinnvolle (reproduzierbare) Beschreibung der Welt und Konstruktion von Mechanismen mit vorhersagbarem Verhalten möglich. Kleine Abweichungen in der Länge des Fadenpendels haben nur kleine Änderungen in der Schwingungsdauer zur Folge, so annähernd gleiche Fadenlängen auch annähernd die selbe Zeit messen. Funktionen, bei denen die Änderungen der Funktionswerte so kontrollierbar sind, heißen *stetige Funktionen.*

DEFINITION 4.1.1. *Eine Funktion $f\colon D \to \mathbb{R}$ heißt* stetig im Punkt $x_0 \in D$, *wenn für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so dass für alle $x \in D$ mit $|x - x_0| < \delta$ gilt $|f(x) - f(x_0)| < \epsilon$. Die Funktion $f\colon D \to \mathbb{R}$ heißt* stetig auf D, *wenn sie stetig in jedem Punkt $x_0 \in D$ ist.*

BEMERKUNG 4.1.2. *Eine geometrische Interpretation der Definition der Stetigkeit ist in Abbildung 12 gegeben: Stetigkeit der Funktion f am Punkt x_0 bedeutet, dass für jeden vorgegebenen ϵ-Streifen um $f(x_0)$ ein δ-Streifen um x_0 existiert, so dass die Funktionswerte von f im vorgegebenen ϵ-Streifen um $f(x_0)$ liegen, wenn die Argumente im δ-Streifen um x_0 liegen.*

BEMERKUNG 4.1.3. *Ein Beispiel für Unstetigkeit ist in Abbildung 13 skizziert: Ein Wagen rollt die schiefe Ebene hinunter, auf der nachfolgenden Geraden*

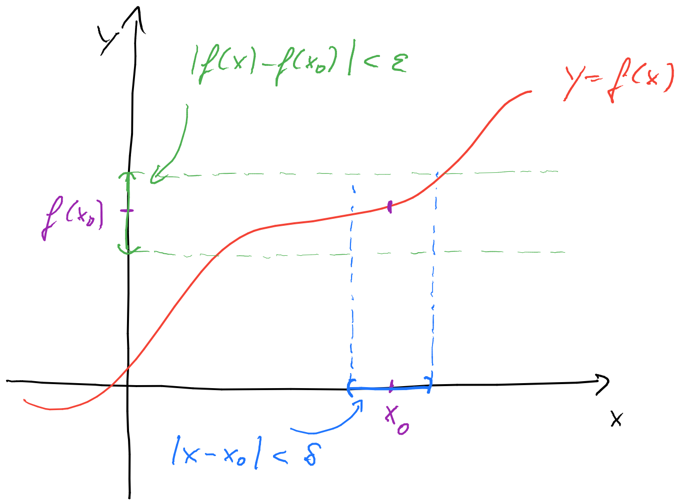


ABBILDUNG 12. Geometrische Interpretation der Stetigkeit: Funktionswerte liegen im ϵ-Streifen um $f(x_0)$, wenn Argumente im δ-Streifen um x_0 liegen.

gibt es ein Hindernis. Hat der Wagen genügend Schwung, kann er das Hindernis überwinden und rollt mit dem zusätzlichen Schwung vom Hindernis noch ein gutes Stück weiter. Hat der Wagen nicht genügend Schwung, kommt er nicht über den Berg und rollt wieder zurück. An der ersten schiefen Ebene gibt es jetzt einen Punkt, so dass der Wagen beim Start unterhalb dieses Punkts zurückrollt und beim Start oberhalb dieses Punkts das Hindernis überwindet. Damit gibt es Startbedingungen, die beliebig nah aneinanderliegen, aber immer zu signifikant unterschiedlichen Ergebnissen führen.

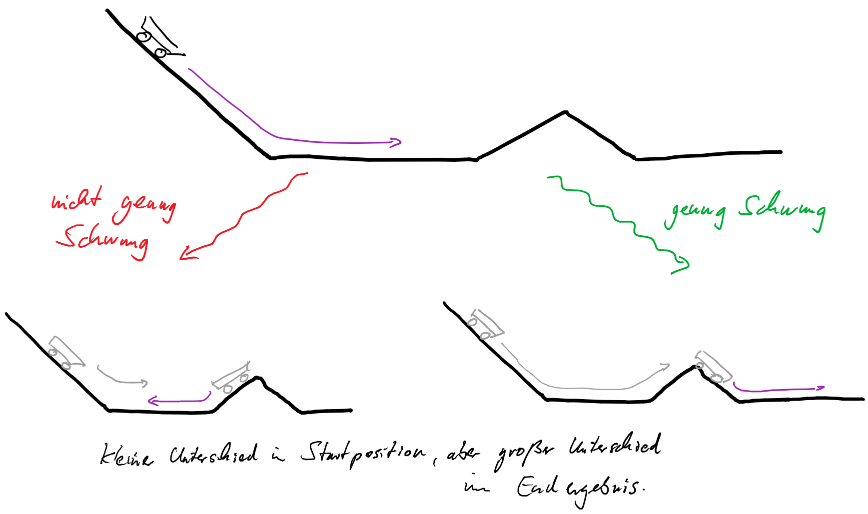


ABBILDUNG 13. Unstetigkeit: beliebig kleine Unterschiede zwischen Startbedingungen führen zu signifikant unterschiedlichen Ergebnissen.

BEISPIEL 4.1.4. *(1) Die Funktion $f(x) = x^2$ ist stetig auf $\mathbb{R}$. Für $x_0 \in \mathbb{R}$ wollen wir $|x^2 - x_0^2| < \epsilon$ erreichen. Mit der Multiplikativität des Betrags ist $|x^2 - x_0^2| = |x + x_0| \cdot |x - x_0|$. Mit Dreiecksungleichung ist $|x + x_0| \leq |x - x_0| + 2|x_0|$. Mit $\delta = \min\left(1, \frac{\epsilon}{2|x_0|+1}\right)$ haben wir dann*

$$|x + x_0| \leq |x - x_0| + 2|x_0| \leq 1 + 2|x_0|, \qquad |x - x_0| < \frac{\epsilon}{1 + 2|x_0|}$$

für alle $|x - x_0| < \delta$. Damit haben wir $|x^2 - x_0^2| < \epsilon$ wie gewünscht.

(2) Für jede natürliche Zahl $k \geq 2$ ist die Wurzelfunktion auf $[0, \infty)$ stetig. Es gilt $|\sqrt[k]{x} - \sqrt[k]{x_0}| \leq \sqrt[k]{|x - x_0|}$. Andernfalls ist im Fall $x > x_0$ mit $\sqrt[k]{x} > \sqrt[k]{x - x_0} + \sqrt[k]{x_0}$ auch

$$x > (\sqrt[k]{x - x_0} + \sqrt[k]{x_0})^k \geq x - x_0 + x_0 + k\,\sqrt[k]{x_0}^{k-1}\,\sqrt[k]{x - x_0} \geq x,$$

und der Fall $x < x_0$ wird analog behandelt. Für gegebenes $\epsilon > 0$ setzen wir $\delta = \epsilon^k$, dann folgt aus $|x - x_0| < \delta$ schon

$$|\sqrt[k]{x} - \sqrt[k]{x_0}| \leq \sqrt[k]{|x - x_0|} < \sqrt[k]{\delta} = \epsilon.$$

(3) Eine Funktion $f\colon D \to \mathbb{R}$ heißt Lipschitz-stetig, *wenn es eine Zahl $L \in \mathbb{R}$ gibt, so dass für alle $x, y \in D$ gilt $|f(x) - f(y)| \leq L|x - y|$. Lipschitz-stetige Funktionen sind stetig (mit $\delta = \frac{\epsilon}{L}$ für $L \neq 0$).*

□

DEFINITION 4.1.5. *Eine Funktion $f\colon D \to \mathbb{R}$ heißt* gleichmäßig stetig, *wenn für alle $\epsilon > 0$ ein $\delta > 0$ existiert, so dass für alle $x_1, x_2 \in D$ mit $|x_1 - x_2| < \delta$ gilt $|f(x_1) - f(x_2)| < \epsilon$.*

BEMERKUNG 4.1.6. *Bei der Definition von Stetigkeit kann δ vom Punkt $x_0 \in D$ abhängen. Bei der Definition von gleichmäßiger Stetigkeit ist δ unabhängig von $x_0 \in D$, ein δ funktioniert für den ganzen Definitionsbereich.*

BEISPIEL 4.1.7. *Die Funktion $f(x) = x^2$ ist auf $\mathbb{R}$ nicht gleichmäßig stetig. Die Differenz $(x+\delta)^2 - x^2 = x^2 + 2\delta x + \delta^2 - x^2 = 2\delta x + \delta^2$. Für vorgegebenes ϵ ist damit die Menge der $x \in \mathbb{R}$, die für ein gewähltes δ die Ungleichung $|2\delta x + \delta^2| < \epsilon$ erfüllen beschränkt. Das heißt, dass wir kein δ finden können, so dass die Ungleichung $|2\delta x + \delta^2| < \epsilon$ für alle $x \in \mathbb{R}$ erfüllt ist.*

Schränken wir allerdings $f(x) = x^2$ auf ein abgeschlossenes Intervall $[a, b]$ ein, dann hat die Menge $\{|x| \mid x \in [a, b]\}$ ein Maximum m. Für vorgegebenes ϵ können wir dann ein δ finden, so dass

$$|2\delta m + \delta^2| \leq 2\delta m + \delta^2 < \epsilon$$

erfüllt ist. Für dieses δ gilt dann auch $|(x + \delta)^2 - x^2| < \epsilon$ für alle $x \in [a, b]$. Die Funktion $f(x) = x^2$ ist also auf dem abgeschlossenen Intervall $[a, b]$ *gleichmäßig stetig.*

□

ÜBUNGSAUFGABE 4.1.8. *Sei $D \subseteq \mathbb{R}$ und $x_0 \in D$. Sei $f\colon D \to \mathbb{R}$ eine in x_0 stetige Funktion mit $f(x_0) \neq 0$. Zeigen Sie, dass ein $\epsilon > 0$ existiert, so dass f in $D \cap (x_0 - \epsilon, x_0 + \epsilon)$ keine Nullstelle hat.*

ÜBUNGSAUFGABE 4.1.9. *Zeigen Sie, dass die Funktion $f\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto \frac{1}{1+|x|}$ gleichmäßig stetig ist.*

ÜBUNGSAUFGABE 4.1.10. *Zeigen Sie, dass die Funktion $f\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto$* $\sin(x^2)$ *nicht gleichmäßig stetig ist.*

ÜBUNGSAUFGABE 4.1.11. *Eine gleichmäßig stetige Funktion $f\colon D \to \mathbb{R}$ mit beschränktem Definitionsbereich ist beschränkt.*

4.2. Grenzwerte von Funktionen und Stetigkeit

Definition 4.2.1. *Sei $D \subseteq \mathbb{R}$ eine Teilmenge. Ein Punkt $x \in \mathbb{R}$ ist ein* Häufungspunkt *von D, wenn in jeder ϵ-Umgebung $(x-\epsilon, x+\epsilon)$ von x unendlich viele Punkte aus D liegen.*

Bemerkung 4.2.2. *Ein Häufungspunkt von D muss nicht notwendigerweise in D liegen. Zum Beispiel sind die Randpunkte a, b eines offenen Intervalls (a, b) Häufungspunkte.*

Definition 4.2.3 (Grenzwerte von Funktionen). *Sei $f\colon D \to \mathbb{R}$ eine Funktion auf $D \subseteq \mathbb{R}$ und $a \in \mathbb{R}$ ein Häufungspunkt von D. Wir definieren*

$$\lim_{x\to a} f(x) = c,$$

wenn für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in D$ und $\lim_{n\to\infty} x_n = a$ gilt $\lim_{n\to\infty} f(x_n) = c$.

Definition 4.2.4 (links- und rechtsseitige Grenzwerte). *Sei $D \subseteq \mathbb{R}$ eine Funktion und $a \in \mathbb{R}$ ein Häufungspunkt von D. Wir definieren*

$$\lim_{x\to a^+} f(x) = c \qquad \textit{bzw.} \qquad \lim_{x\to a^-} f(x) = c$$

wenn für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in D$, $x_n > a$ (bzw. $x_n < a$) für alle x_n und $\lim_{n\to\infty} x_n = a$ gilt $\lim_{n\to\infty} f(x_n) = c$.

Beispiel 4.2.5.
1. *Wir betrachten $f(x) = x^2$ und $a \in \mathbb{R}$. Sei $(x_n)_n$ eine Folge mit $\lim_{n\to\infty} x_n = a$. Dann ist $\lim_{n\to\infty} x_n^2 = a^2$ nach den Grenzwertrechenregeln aus Proposition 2.2.1.*
2. *Wir berechnen $\lim_{x\to 0} \frac{\sin x}{x}$ mit einer geometrischen Überlegung (d.h. mit der Definition von Sinus als Winkelfunktion im rechtwinkligen Dreieck). Der Definitionsbereich von $f(x) = \frac{\sin x}{x}$ ist $\mathbb{R}\setminus\{0\}$, und hat insbesondere 0 als Häufungspunkt. Am Einheitskreis betrachten wir den Kreissektor mit Winkel x, $|x| < \frac{\pi}{2}$, dessen Fläche wir durch zwei rechtwinklige Dreiecke nach oben bzw. unten abschätzen können.*

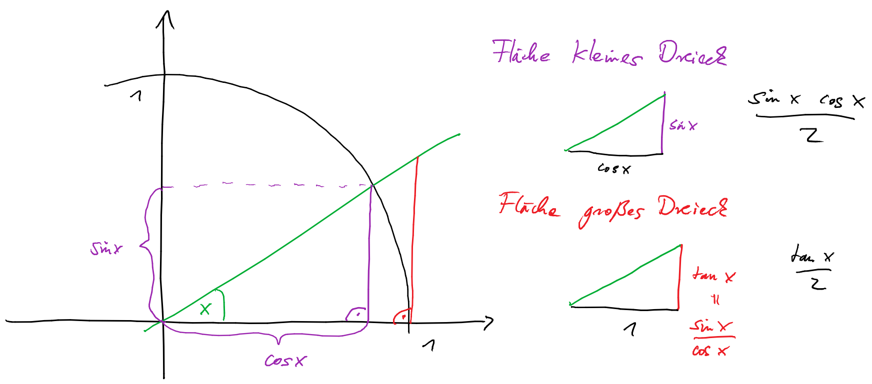


Abbildung 14. Abschätzung für $\lim_{x\to 0} \frac{\sin x}{x}$

Der Kreissektor hat Fläche $\frac{x}{2}$, da der Einheitskreis die Fläche π hat und wir x im Bogenmaß messen. Die Fläche des kleineren Dreiecks ist $\frac{\sin x \cdot \cos x}{2}$, da die beiden Katheten des rechtwinkligen Dreiecks die Länge $\sin x$ bzw. $\cos x$ haben. Die Fläche des größeren Dreiecks ist

$$\frac{\tan x}{2} = \frac{\sin x}{2\cos x},$$

da die Katheten des Dreiecks die Länge 1 bzw. $\frac{\sin x}{\cos x}$ *haben. Wir haben damit für* $|x| < \pi$ *die Abschätzung*

$$\frac{\sin x \cdot \cos x}{2} \leq \frac{x}{2} \leq \frac{\sin x}{2 \cos x}.$$

Teilen wir durch $\frac{\sin x}{2}$ $(\neq 0$ *für* $x \neq 0)$ *und invertieren* $(\cos x \neq 0$ *für* $|x| < \frac{\pi}{2})$ *ergibt sich*

$$\cos x \leq \frac{\sin x}{x} \leq \frac{1}{\cos x}.$$

Für eine Nullfolge $(x_k)_{k\in\mathbb{N}}$ *mit* $x_k \neq 0$ *für alle* k *haben wir dann*

$$\lim_{k\to\infty} \cos x_k = \lim_{k\to\infty} \frac{1}{\cos x_k} = \lim_{k\to\infty} \frac{\sin x_k}{x_k} = 1$$

(3) Wir betrachten die Funktion

$$f : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

Dann ist $\lim_{x\to 0^+} f(x) = 1$ *und* $\lim_{x\to 0^-} f(x) = 0$.

□

PROPOSITION 4.2.6. *Sei* $D \subseteq \mathbb{R}$ *und* $f : D \to \mathbb{R}$ *eine Funktion, und* $a \in D$ *ein Häufungspunkt von* D*. Die beiden folgenden Bedingungen sind äquivalent:*

(1) $\forall \epsilon > 0\ \exists \delta > 0\ \forall x \in D : (|x - a| < \delta) \Rightarrow (|f(x) - f(a)| < \epsilon)$

(2) $\lim_{x\to a} f(x) = f(a)$.

Stetigkeit und Folgenstetigkeit sind äquivalent.

BEWEIS. "⇒": Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in D mit $\lim_{n\to\infty} x_n = a$. Wir wollen $\lim_{n\to\infty} f(x_n) = f(a)$ zeigen, d.h. für $\epsilon > 0$ suchen wir ein N, so dass für $n \geq N$ gilt $|f(x_n) - f(a)| < \epsilon$. Nach Voraussetzung (1) existiert ein $\delta > 0$, so dass für alle $x \in D$ mit $|x - a| < \delta$ gilt $|f(x) - f(a)| < \epsilon$. Das bedeutet, dass für alle n mit $|x_n - a| < \delta$ die gewünschte Abschätzung $|f(x_n) - f(a)| < \epsilon$ gilt. Da aber $\lim_{n\to\infty} x_n = a$ gilt, existiert ein N, so dass für alle $n \geq N$ gilt $|x_n - a| < \delta$. Damit haben wir die Behauptung $\lim_{n\to\infty} f(x_n) = f(a)$ gezeigt.

"⇐": Wir beweisen die Aussage durch Widerspruch. Wir nehmen an, dass (1) nicht gilt, und wollen zeigen, dass eine Folge $(x_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} x_n = a$ existiert, so dass $f(x_n)$ nicht gegen $f(a)$ konvergiert. Die Verneinung von (1) ist

$$\exists \epsilon > 0\ \forall \delta > 0\ \exists x \in D : (|x - a| < \delta) \wedge (|f(x) - f(a)| \geq \epsilon).$$

Für eine natürliche Zahl n nehmen wir dieses ϵ und erhalten für $\delta = \frac{1}{n}$, dass ein $x_n \in D$ existiert, so dass $|x_n - a| < \delta$ und $|f(x_n) - f(a)| \geq \epsilon$ gilt. Die so konstruierte Folge konvergiert wegen $|x_n - a| < \frac{1}{n}$ gegen a, aber wegen $|f(x_n) - f(a)| \geq \epsilon$ konvergieren die Funktionswerte $f(x_n)$ nicht gegen $f(a)$. □

ÜBUNGSAUFGABE 4.2.7. *Wir betrachten für* $n \in \mathbb{N}$ *die Funktionen*

$$g_n : \mathbb{R} \to \mathbb{R} : x \mapsto \frac{nx}{1 + |nx|}.$$

Zeigen Sie, dass alle Funktionen g_n *stetig sind. Für welche* $x \in \mathbb{R}$ *ist die Funktion* $g(x) := \lim_{n\to\infty} g_n(x)$ *definiert bzw. stetig?*

DEFINITION 4.2.8. *Sei* $f : D \to \mathbb{R}$ *eine stetige Funktion und* $a \in \mathbb{R} \setminus D$*. Wir sagen, dass* f *eine* stetige Fortsetzung *in den Punkt* a *hat, wenn eine stetige Funktion* $\bar{f} : D \cup \{a\} \to \mathbb{R}$ *mit* $\bar{f}|_D = f$ *existiert.*

BEMERKUNG 4.2.9. *Wenn a kein Häufungspunkt von D ist, dann ist für alle c die Funktion*

$$\bar{f}(x) = \begin{cases} c & x = a \\ f(x) & x \in D \end{cases}$$

stetig in a.

PROPOSITION 4.2.10. *Sei $f\colon D \to \mathbb{R}$ eine stetige Funktion und $a \in \mathbb{R} \setminus D$ ein Häufungspunkt von D. Dann existiert höchstens eine stetige Fortsetzung $\bar{f}\colon D \cup \{a\} \to \mathbb{R}$.*

BEWEIS. Stetigkeit in a bedeutet, dass für alle Folgen $(x_n)_n$ mit $x_n \in D$ und $\lim_{n\to\infty} x_n = a$ gilt $\lim_{n\to\infty} f(x_n) = c$ für ein $c \in \mathbb{R}$. Wenn dieser Grenzwert für alle solchen Folgen existiert (und immer gleich c ist), dann ist $\bar{f}(a) = c$ die eindeutige stetige Fortsetzung von f auf $D \cup \{a\}$. □

BEISPIEL 4.2.11. *Die Funktion $f(x) = 1$ auf dem Definitionsbereich $\mathbb{R} \setminus \{0\}$ hat eine stetige Fortsetzung auf $\mathbb{R}$. Die Funktion*

$$f\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}\colon x \mapsto \begin{cases} 0 & x < 0 \\ 1 & x > 0 \end{cases}$$

hat keine stetige Fortsetzung auf $\mathbb{R}$. □

DEFINITION 4.2.12. *Sei $f\colon D \to \mathbb{R}$ eine Funktion, deren Definitionsbereich $D \subseteq \mathbb{R}$ nicht nach oben beschränkt ist. Wir definieren*

$$\lim_{x\to\infty} f(x) = c,$$

wenn für jedes $\epsilon > 0$ eine Zahl N existiert, so dass $|f(x) - c| < \epsilon$ für alle $x \in D$ mit $x > N$ gilt. Analog definieren wir $\lim_{x\to-\infty} f(x)$.

DEFINITION 4.2.13. *Sei $f\colon D \to \mathbb{R}$ eine Funktion und $a \in \mathbb{R}$ ein Häufungspunkt von D. Wenn für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} x_n = a$ die Folge $f(x_n)$ bestimmt gegen ∞ divergiert, dann schreiben wir*

$$\lim_{x\to a} f(x) = \infty.$$

Analog für die bestimmte Divergenz gegen $-\infty$; ebenso für einseitige Grenzwerte $\lim_{x\to a^\pm} f(x)$ bzw. Grenzwerte $\lim_{x\to\pm\infty} f(x)$.

4.3. Rechenregeln für Stetigkeit

PROPOSITION 4.3.1. *Seien $f, g\colon D \to \mathbb{R}$ stetig in $a \in D$. Dann sind auch $f + g$ und $f \cdot g$ stetig in a. Ist $g(a) \neq 0$, dann gibt es ein $\epsilon > 0$, so dass $\frac{f}{g}$ auf $D \cap (a-\epsilon, a+\epsilon)$ definiert und stetig ist.*

BEWEIS. Wir benutzen die Folgenstetigkeit: für eine Folge (x_n) mit $x_n \in D$ und $\lim_{n\to\infty} x_n = a$ haben wir nach Voraussetzung $\lim_{n\to\infty} f(x_n) = f(a)$ und $\lim_{n\to\infty} g(x_n) = g(a)$. Mit den Rechenregeln für Folgengrenzwerte, s. Proposition 2.2.1, gilt

$$\lim_{n\to\infty} (f+g)(x_n) = (f+g)(a), \qquad \lim_{n\to\infty} (f \cdot g)(x_n) = (f \cdot g)(a),$$

also sind $f + g$ und $f \cdot g$ stetig.

Nach Übungsaufgabe 4.1.8 existiert ein $\epsilon > 0$, so dass g auf $D \cap (a - \epsilon, a + \epsilon)$ keine Nullstelle hat. Damit ist $\frac{f}{g}$ definiert und mit Proposition 2.2.1 haben wir die Stetigkeit. □

BEISPIEL 4.3.2. *Polynome und rationale Funktionen sind stetig auf dem gesamten Definitionsbereich.* □

PROPOSITION 4.3.3. *Seien $D, D' \subseteq \mathbb{R}$ und $f\colon D \to \mathbb{R}$ und $g\colon D' \to \mathbb{R}$ Funktionen mit $f(D) \subseteq D'$. Wenn f stetig in a und g stetig in $f(a)$ ist, dann ist $g \circ f\colon D \to \mathbb{R}$ stetig in a.*

BEWEIS. Für eine Folge (x_n) mit $x_n \in D$ und $\lim_{n\to\infty} x_n = a$ ist $(f(x_n))_n$ eine Folge in D' mit $\lim_{n\to\infty} f(x_n) = f(a)$, da f stetig ist. Da g stetig ist, ist $\lim_{n\to\infty} g(f(x_n)) = g(f(a))$, also ist $g \circ f$ stetig. □

PROPOSITION 4.3.4. *Sei $(f_n\colon D \to \mathbb{R})_{n\in\mathbb{N}}$ eine gleichmäßig konvergente Folge stetiger Funktionen. Dann ist auch die Grenzfunktion $\lim_{n\to\infty} f_n$ stetig.*

BEWEIS. Sei $a \in D$ und $\epsilon > 0$. Wir bezeichnen die Grenzfunktion mit $f = \lim_{n\to\infty} f_n$ und wollen zeigen, dass ein $\delta > 0$ existiert, so dass für alle $x \in D$ mit $|x - a| < \delta$ gilt $|f(x) - f(a)| < \epsilon$. Für diese Abschätzung benutzen wir die Funktionen f_n der Folge als Bezugspunkte:

$$|f(x) - f(a)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)|.$$

Um $|f(x)-f(a)| < \epsilon$ zu erhalten, reicht es zu zeigen, dass alle Terme auf der rechten Seite $< \frac{\epsilon}{3}$ sind. Aus der gleichmäßigen Konvergenz von $(f_n)_{n\in\mathbb{N}}$ folgt, dass es ein n gibt, so dass $|f(x) - f_n(x)| < \frac{\epsilon}{3}$ für alle $x \in D$ gilt. Aus der Stetigkeit von f_n folgt, dass es ein $\delta > 0$ gibt, so dass für alle $x \in D$ mit $|x-a| < \delta$ gilt $|f_n(x)-f_n(a)| < \frac{\epsilon}{3}$. Mit diesem δ gilt dann für alle $x \in D$ mit $|x - a| < \delta$ auch $|f(x) - f(a)| < \epsilon$, und wir haben die Stetigkeit von f gezeigt. □

BEISPIEL 4.3.5. *Eine Potenzreihe $\sum_{n=0}^{\infty} a_n z^n$ stellt auf abgeschlossenen Intervallen im Konvergenzkreis eine stetige Funktion dar. Dies gilt insbesondere für Exponentialfunktion, Logarithmus etc.* □

Analog übertragen sich Rechenregeln zum Berechnen von Funktionengrenzwerten.

BEISPIEL 4.3.6. (1) *Für $f\colon \mathbb{R}\setminus\{0\} \to \mathbb{R}\colon x \mapsto \frac{1}{x}$ ist*

$$\lim_{x\to 0^+} f(x) = +\infty \qquad \text{und} \qquad \lim_{x\to 0^-} f(x) = -\infty.$$

(2) *Für ein Polynom $p(x) = \sum_{n=0}^{d} a_n x^n$ mit $a_d \neq 0$ haben wir die folgenden Fälle:*
- *Wenn n gerade ist, dann ist $\lim_{x\to\pm\infty} p(x) = \infty$ für $a_d > 0$ und $\lim_{x\to\pm\infty} p(x) = -\infty$ für $a_d < 0$.*
- *Wenn n ungerade ist und $a_d > 0$, dann ist $\lim_{x\to\infty} p(x) = \infty$ und $\lim_{x\to-\infty} p(x) = -\infty$. Für den Fall $a_d > 0$ drehen sich die Vorzeichen um.*

(3) *Für eine rationale Funktion $f(x) = \frac{p(x)}{q(x)}$ mit $p(x) = \sum_{n=0}^{\deg p} a_n x^n$ und $q(x) = \sum_{m=0}^{\deg q} b_m x^n \neq 0$ haben wir folgende Fälle:*
- *Wenn $\deg p(x) > \deg q(x)$, dann verhält sich $\lim_{x\to\pm\infty} f(x)$ wie das Polynom $\frac{a_{\deg p}}{b_{\deg q}} x^{\deg p - \deg q}$.*
- *Wenn $\deg p(x) = \deg q(x) = d$, dann ist*

$$\lim_{x\to\pm\infty} = \frac{a_d}{b_d}$$

- *Wenn $\deg p(x) < \deg q(x)$, dann ist*

$$\lim_{x\to\pm\infty} = 0.$$

(4) Allgemeiner können wir das asymptotische Verhalten von Funktionen f, g (auf entsprechend unbeschränkten Definitionsbereichen) mittels Grenzwerten von der Form

$$\lim_{x \to \pm\infty} \frac{f(x)}{g(x)}$$

vergleichen.

□

Übungsaufgabe 4.3.7. *Bestimmen Sie die folgenden Funktionengrenzwerte:*

(a) $\lim_{x \to 1} \frac{x^3-1}{(x-1)^3}$

(b) $\lim_{x \to 1} \frac{x^3+x^2-x-1}{x-1}$

(c) $\lim_{x \to 0} \frac{1-\sqrt{1-x^2}}{x^2}$

(d) $\lim_{x \to 0} \frac{x^2}{|x|}$

Übungsaufgabe 4.3.8. *Untersuchen Sie, ob*

$$\lim_{x \to 0} \frac{1 \pm \sqrt{1-4x}}{2x}.$$

existiert. Geben Sie den Grenzwert an, falls er existiert.

Hinweis: *Führen Sie eine Grenzwertbetrachtung für den Kehrwert* $\frac{2x}{1 \pm \sqrt{1-4x}}$ *durch.*

4.4. Stetige Funktionen auf abgeschlossenen Intervallen

Stetigkeit hat für Funktionen auf abgeschlossenen Intervallen $[a, b]$ einige starke Konsequenzen, die wir im Folgenden auflisten wollen.

Satz 4.4.1 (Zwischenwertsatz). *Sei $f\colon [a, b] \to \mathbb{R}$ eine stetige Funktion. Für jedes y mit $f(a) \le y \le f(b)$ existiert ein $x_0 \in [a, b]$ mit $f(x_0) = y$.*

Beweis. Wir geben eine Intervallschachtelung an, die x_0 definiert. Die Intervallschachtelung $I_n = [a_n, b_n]$ wird wieder rekursiv konstruiert und soll $f(a_n) \le y \le f(b_n)$ erfüllen. Der Rekursionsanfang ist durch $a_1 = a$ und $b_1 = b$ gegeben. Im Rekursionsschritt halbieren wir das gegebene Intervall $I_k = [a_k, b_k]$ durch den Mittelpunkt $m = \frac{a_k+b_k}{2}$ und setzen

$$[a_{k+1}, b_{k+1}] = \begin{cases} [a_{k+1}, m] & f(m) \ge y \\ [m, b_{k+1}] & f(m) < y \end{cases}$$

Dann gilt $f(a_{k+1}) \le y \le f(b_{k+1})$ nach Konstruktion, und wir erhalten eine Intervallschachtelung.

Sei x_0 die Zahl, die in allen Intervallen enthalten ist. Es bleibt noch zu zeigen, dass $f(x_0) = y$ gilt. Die Folgen $(a_k)_{k \ge 1}$ und $(b_k)_{k \ge 1}$ konvergieren nach Konstruktion gegen x_0. Aus der Stetigkeit von f folgt damit auch $\lim_{k \to \infty} f(a_k) = f(x_0) = \lim_{k \to \infty} f(b_k)$. Nach Konstruktion gilt aber auch $f(a_k) \le y \le f(b_k)$, und wir erhalten $f(x_0) = y$. □

Beispiel 4.4.2. *Wir können mit dem Zwischenwertsatz noch einmal die Existenz von Wurzeln reeller Zahlen beweisen. Dazu betrachten wir für $\alpha \in \mathbb{R}_+$ und eine natürliche Zahl $n \ge 2$ das Polynom $p(x) = x^n - \alpha$. Da $p(0) = \alpha < 0$ ist und mit der Bernoulli-Ungleichung $p(1 + \alpha) = (1 + \alpha)^n - \alpha \ge 1 + n\alpha - \alpha > 0$ ist, muss das Polynom $p(x)$ nach dem Zwischenwertsatz 4.4.1 eine positive Nullstelle y haben, für die dann $p(y) = y^n - \alpha = 0$ gilt. Damit haben wir die Existenz einer n-ten Wurzel aus x gezeigt.*

□

KOROLLAR 4.4.3. *Sei $I \subseteq \mathbb{R}$ ein Intervall (offen, abgeschlossen, evtl. mit Intervallgrenzen $\pm\infty$) und $f\colon I \to \mathbb{R}$ eine stetige Funktion. Dann ist auch $f(I) \subseteq \mathbb{R}$ ein Intervall.*

BEWEIS. Wir definieren $B = \sup f(I)$ und $A = \inf f(I)$ (wobei wir $A = -\infty$ bzw. $B = +\infty$ setzen, falls $f(I)$ nicht nach unten bzw. oben beschränkt ist). Wir zeigen, dass $(A, B) \subseteq f(I)$. Für y mit $A < y < B$ existieren nach Definition $a, b \in I$ mit $f(a) < y < f(b)$. Nach dem Zwischenwertsatz 4.4.1 existiert ein x mit $f(x) = y$, also ist $y \in f(I)$. Damit ist $(A, B) \subseteq f(I)$, und $f(I)$ muss gleich einem der Intervalle (A, B), $[A, B]$, $(A, B]$ oder $[A, B)$ sein. $\square$

ÜBUNGSAUFGABE 4.4.4. *Jedes reelle Polynom hat eine reelle Nullstelle.*

ÜBUNGSAUFGABE 4.4.5. *Sei $n > 1$ eine natürliche Zahl. Zeigen Sie, dass es keine stetige reelle Funktion auf $[0, 1]$ gibt, die jeden ihrer Werte genau n-mal annimmt.*

ÜBUNGSAUFGABE 4.4.6. *Zeigen Sie, dass eine stetige Funktion $f\colon I \to \mathbb{R}$ auf einem Intervall genau dann injektiv ist, wenn sie streng monoton ist.*

SATZ 4.4.7. *Eine stetige Funktion $f\colon [a, b] \to \mathbb{R}$ auf einem abgeschlossenen Intervall ist beschränkt und hat im Intervall ein Minimum und ein Maximum.*

BEWEIS. Wir beweisen die Aussage für das Maximum, die Aussage für das Minimum folgt durch Anwendung auf $-f$. Sei

$$A = \sup\{f(x) \mid x \in [a, b]\}.$$

Dann existiert eine Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in [a, b]$, so dass $\lim_{n\to\infty} f(x_n) = A$. (Falls die Menge $\{f(x) \mid x \in [a, b]\}$ nicht beschränkt ist, haben wir eine Folge, die bestimmt gegen $+\infty$ divergiert.) Da die Folge $(x_n)_n$ beschränkt ist, existiert nach Satz 2.3.12 (Bolzano–Weierstraß) eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ mit $\lim_{k\to\infty} x_{n_k} = p \in [a, b]$. Aus der Stetigkeit von f folgt $f(p) = \lim_{k\to\infty} f(x_{n_k}) = A$. Insbesondere ist A definiert, damit ist f nach oben beschränkt und nimmt in p das Maximum an. $\square$

BEMERKUNG 4.4.8. *Es gibt topologischere Formulierungen dieser Aussage, die den Begriff von kompakten Mengen benutzen (z.B. im Königsberger). Die Konzepte von Topologie, offenen, abgeschlossenen und kompakten Mengen werden am Anfang der Analysis 2 diskutiert, wenn wir Funktionen auf Teilmengen des $\mathbb{R}^n$ betrachten.*

PROPOSITION 4.4.9. *Sei $f\colon [a, b] \to \mathbb{R}$ eine auf einem abgeschlossenen Intervall stetige Funktion. Dann ist f auch gleichmäßig stetig.*

BEWEIS. Wir nehmen an, dass f nicht gleichmäßig stetig ist. Dann existiert ein $\epsilon > 0$, so dass wir für jedes $n \geq 1$ zwei Punkte $x_n, y_n \in [a, b]$ mit $|x_n - y_n| < \frac{1}{n}$ und $|f(x_n) - f(y_n)| \geq \epsilon$ finden können. Nach Satz 2.3.12 (Bolzano–Weierstraß) existiert für die beschränkte Folge $(x_n)_{n\in\mathbb{N}}$ eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ mit Grenzwert $\lim_{k\to\infty} x_{n_k} = p \in [a, b]$. Mit $|x_{n_k} - y_{n_k}| < \frac{1}{n_k}$ konvergiert auch $(y_{n_k})_{k\in\mathbb{N}}$ gegen p. Da f stetig ist, folgt

$$\lim_{k\to\infty} (f(x_{n_k}) - f(y_{n_k})) = f(p) - f(p) = 0.$$

Dies widerspricht $|f(x_n) - f(y_n)| \geq \epsilon$ aus der Konstruktion. Damit muss f gleichmäßig stetig sein. $\square$

ÜBUNGSAUFGABE 4.4.10. *Eine stetige Funktion $f\colon (a, b) \to \mathbb{R}$ besitzt genau dann eine stetige Fortsetzung auf $[a, b]$, wenn sie auf (a, b) gleichmäßig stetig ist.*

PROPOSITION 4.4.11. *Sei $f\colon [a, b] \to \mathbb{R}$ eine stetige Funktion. Dann gibt es zu jedem $\epsilon > 0$ stückweise konstante Funktionen $\phi, \psi\colon [a, b] \to \mathbb{R}$, so dass gilt*

(1) $\phi(x) \leq f(x) \leq \psi(x)$ *für alle* $x \in [a, b]$
(2) $|\phi(x) - \psi(x)| \leq \epsilon$ *für alle* $x \in [a, b]$

Beweis. Nach Proposition 4.4.9 ist f gleichmäßig stetig. Es existiert also für jedes $\epsilon > 0$ ein $\delta > 0$, so dass für alle $x, y \in [a, b]$ mit $|x-y| < \delta$ gilt $|f(x)-f(y)| < \epsilon$. Wir wählen n so, dass $\frac{b-a}{n} < \delta$ und setzen $t_k = a + k\frac{b-a}{n}$. Damit unterteilen wir das Intervall $[a, b]$ durch $a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$ in Teilintervalle $[t_i, t_{i+1}]$ mit Länge $t_i - t_{i-1} < \delta$. Für $1 \leq k \leq n$ setzen wir

$$c_k = \sup\{f(x) \mid x \in [t_{k-1}, t_k]\}, \qquad d_k = \inf\{f(x) \mid x \in [t_{k-1}, t_k]\}.$$

Nach dem Satz über Existenz von Minimum und Maximum 4.4.7 werden c_k und d_k in $[t_{k-1}, t_k]$ angenommen. Damit gilt $|c_k - d_k| < \epsilon$ für alle k. Damit können wir die stückweise konstanten Funktionen definieren, indem wir setzen:

$$\phi(a) = \psi(a) = f(a), \qquad \phi(x) = d_k, \psi(x) = c_k \text{ für } x \in [t_{k-1}, t_k].$$

□

4.5. Umkehrsatz

Definition 4.5.1. *Sei* $D \subseteq \mathbb{R}$ *und* $f \colon D \to \mathbb{R}$. *Die Funktion* f *heißt* (streng) monoton wachsend, *wenn für alle* $x < y$ *in* D *gilt* $f(x) \leq f(y)$ *(bzw.* $f(x) < f(y)$*). Die Funktion* f *heißt* (streng) monoton fallend, *wenn für alle* $x < y$ *in* D *gilt* $f(x) \geq f(y)$ *(bzw.* $f(x) > f(y)$*).*

Satz 4.5.2. *Sei* $f \colon I \to \mathbb{R}$ *auf einem Intervall streng monoton wachsend (bzw. fallend) und stetig. Dann existiert eine Umkehrfunktion* $f^{-1} \colon f(I) \to \mathbb{R}$*, die stetig und streng monoton wachsend (bzw. fallend) ist.*

Beweis. Wir beweisen den Fall monoton wachsend, der Fall monoton fallend ist komplett analog. Nach Korollar 4.4.3 ist $f(I)$ ein Intervall. Da f streng monoton ist, ist f injektiv. Damit ist die Abbildung $f \colon I \to f(I)$ bijektiv, d.h. die Umkehrabbildung $f^{-1} \colon f(I) \to I \subseteq \mathbb{R}$ existiert (und bildet $f(x)$ auf x ab). Die Umkehrabbildung ist wieder streng monoton wachsend.

Wir müssen noch die Stetigkeit von f^{-1} zeigen. Sei $b \in f(I)$ gegeben, und sei $a = f^{-1}(b)$. Wir zeigen, dass f^{-1} im Punkt b stetig ist. Nehmen wir zuerst an, dass b kein Randpunkt von $f(I)$ ist. Dann ist auch a kein Randpunkt von I. Für $\epsilon > 0$ mit $[a - \epsilon, a + \epsilon] \subseteq D$ betrachten wir $b_1 = f(a - \epsilon)$ und $b_2 = f(a + \epsilon)$. Dann ist $b_1 < b < b_2$ und f bildet $[a - \epsilon, a + \epsilon]$ bijektiv auf $[b_1, b_2]$ ab. Wir setzen $\delta = \min(b - b_1, b_2 - b)$ und sehen, dass f^{-1} das Intervall $(b - \delta, b + \delta)$ in das Intervall $(a - \epsilon, a + \epsilon)$ abbildet. Damit ist f^{-1} stetig in b. Der Fall, dass b ein Randpunkt ist, wird analog (mit Intervallen $[a - \epsilon, a]$ bzw. $[a, a + \epsilon]$) bewiesen. □

Beispiel 4.5.3. *Aus der Stetigkeit (und strengen Monotonie) der Potenzfunktion* $f \colon \mathbb{R}_+ \to \mathbb{R} \colon x \mapsto x^k$ *folgt mit dem Umkehrsatz 4.5.2 auch die Stetigkeit der Wurzelfunktion* $x \mapsto \sqrt[k]{x}$*. Wenn* k *ungerade ist, ist die Potenzfunktion auf ganz* $\mathbb{R}$ *streng monoton (und stetig) und die Wurzelfunktion* $x \to \sqrt[k]{x}$ *kann auf ganz* $\mathbb{R}$ *definiert werden.* □

4.6. Stetigkeit komplexwertiger Funktionen

Wir diskutieren kurz die Stetigkeit komplexwertiger Funktionen. Definition 4.1.1 funktioniert genausogut im Komplexen: eine Funktion $f \colon D \to \mathbb{C}$ mit $D \subseteq \mathbb{C}$ heißt stetig am Punkt z_0, wenn für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so dass für alle $z \in D$ mit $\|z - z_0\| < \delta$ gilt $\|f(z) - f(z_0)\| < \epsilon$. Dies ist wieder äquivalent zur Folgenstetigkeit, d.h. eine Funktion $f \colon D \to \mathbb{C}$ ist genau dann stetig am Punkt a, wenn für jede Folge $(z_n)_n$ mit $z_n \in D$ und $\lim_{n\to\infty} z_n = a$ gilt $\lim_{n\to\infty} f(z_n) = f(a)$.

Die Rechenregeln für Addition, Multiplikation und Division stetiger Funktionen übertragen sich, ebenso wie die Aussage, dass die Grenzfunktion einer gleichmäßig konvergenten Funktionenfolge von stetigen Funktionen wieder stetig ist.

Im Unterschied zu Funktionen auf Definitionsbereichen $D \subseteq \mathbb{R}$ gibt es keine einseitigen Grenzwerte mehr. Ebenso können wir Zwischenwertsatz und Existenz von Minima bzw. Maxima für komplexwertige Funktionen nicht mehr formulieren. Eine reellwertige Funktion $f\colon D \to \mathbb{R}$ auf einer beschränkten und unter Grenzwerten von Folgen abgeschlossenen Teilmenge $D \subseteq \mathbb{C}$ hat in D immer ein Minimum und ein Maximum (das gilt zum Beispiel für den Betrag einer komplexwertigen Funktion), das beweisen wir hier aber nicht.

KAPITEL 5

Exponentialfunktion, Logarithmus und Winkelfunktionen

5.1. Exponentialfunktion und Logarithmus

Bei vielen natürlichen Wachstumsprozessen (Bakterienwachstum bei ausreichend Ressourcen, Infektionsfälle zum Beginn einer Epidemie, weniger natürlich Wachstum verzinster Darlehen/Schulden,...) bzw. Zerfallsprozessen (Bierschaum, radioaktiver Zerfall zum Beispiel C^{14} bei der Radiokarbondatierung, Temperaturangleichung bei der Wärmeleitung,...) sind die Änderungsraten proportional zum aktuellen Wert: Bakterien vermehren sich proportional zur Anzahl der vorhandenen Bakterien, Zinsen sind ein festgelegter Prozentsatz des aktuellen Darlehens, die Anzahl der zerfallenden Atome in einem festgelegten Zeitintervall ist proportional zur Gesamtzahl der vorhandenen Atome. Solche Prozesse werden annähernd durch

$$f(t + \Delta t) = f(t) + \alpha f(t)\Delta t = (1 + \alpha\Delta t)f(t)$$

beschrieben. Wir können einen solchen Prozess über einen Zeitraum $[0, T]$ modellieren. Dafür unterteilen wir $[0, T]$ in Zeitintervalle

$$[t_0 = 0, t_1], [t_1, t_2], \ldots, [t_{n-1}, t_n = T],$$

für die jeweils $f(t_k) = (1 + \alpha(t_k - t_{k-1}))f(t_{k-1})$ gilt. Haben die Intervalle alle die gleiche Länge $\Delta T = T/n$, erhalten wir dann induktiv für den Zustand zum Zeitpunkt t_k

$$f(t_k) = \left(1 + \alpha\frac{T}{n}\right)^k f(0),$$

insbesondere $f(T) = \left(1 + \alpha\frac{T}{n}\right)^n f(0)$. Die Näherungen werden besser, je feinere Unterteilungen wir wählen – wir fragen uns also, ob der Grenzwert

$$\lim_{n\to\infty}\left(1 + \frac{\alpha T}{n}\right)^n$$

existiert.[1]

Proposition 5.1.1. *Für jede (komplexe) Folge $(w_n)_n$ mit Grenzwert w gilt*

$$\lim_{n\to\infty}\left(1 + \frac{w_n}{n}\right)^n = \sum_{k=0}^{\infty}\frac{w^k}{k!}.$$

Beweis. Für $\epsilon > 0$ existiert ein Index K, so dass für alle $n \geq K$ die beiden folgenden Abschätzungen gelten:

$$\sum_{k=K}^{\infty}\frac{(\|w\| + 1)^k}{k!} < \frac{\epsilon}{3}, \qquad \|w_n\| \leq \|w\| + 1.$$

[1]Historisch tauchen Fragen zur Exponentialfunktion auf, als sich Jakob Bernoulli Gedanken über kontinuierliche Verzinsung macht.

(Die Existenz von K folgt aus der Konvergenz der Exponentialreihe, s. Beispiel 2.6.12, sowie der Konvergenz der Folge (w_n).) Mit dem binomischen Satz und der Dreiecksungleichung haben wir für jedes $n \geq K$

$$\left\| \left(1 + \frac{w_n}{n}\right)^n - \sum_{k=0}^{\infty} \frac{w^k}{k!} \right\| \leq \sum_{k=0}^{K-1} \left\| \binom{n}{k} \frac{w_n^k}{n^k} - \frac{w^k}{k!} \right\| + \sum_{k=K}^{n} \binom{n}{k} \frac{\|w_n\|^k}{n^k} + \sum_{k=K}^{\infty} \frac{\|w\|^k}{k!}.$$

Nach Wahl von K ist die letzte Summe kleiner als $\frac{\epsilon}{3}$. Für die mittlere Summe haben wir

$$\binom{n}{k} \frac{1}{n^k} = \frac{1}{k!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \leq \frac{1}{k!}$$

und erhalten mit der Wahl von K

$$\sum_{k=K}^{n} \binom{n}{k} \frac{\|w_n\|^k}{n^k} < \sum_{k=K}^{n} \frac{(\|w\| + 1)^k}{k!} < \frac{\epsilon}{3}.$$

Die erste Summe konvergiert wegen $\lim_{n\to\infty} \binom{n}{k} \frac{1}{n^k} = \frac{1}{k!}$ und $\lim_{n\to\infty} w_n = w$ für $n \to \infty$ gegen 0. Damit existiert ein $N > K$, so dass die erste Summe auf der rechten Seite für $n > N$ kleiner als $\frac{\epsilon}{3}$ wird. Damit haben wir die Behauptung

$$\left\| \left(1 + \frac{w_n}{n}\right)^n - \sum_{k=0}^{\infty} \frac{w^k}{k!} \right\| < \epsilon.$$

□

Definition 5.1.2. *Wir definieren die komplexe Exponentialfunktion* $\exp\colon \mathbb{C} \to \mathbb{C}$ *durch*

$$\exp(z) := \lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n = \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

Proposition 5.1.3. *Die Exponentialreihe $\sum_{k=0}^{\infty} \frac{z^k}{k!}$ konvergiert auf $\mathbb{C}$ absolut und gleichmäßig.*

Beweis. Nach Proposition 3.2.8 ist der Konvergenzradius

$$\rho = \left(\lim_{n\to\infty} \sqrt[n]{\frac{1}{n!}}\right)^{-1} = \lim_{n\to\infty} \sqrt[n]{n!}.$$

Wir wollen zeigen, dass $\rho = \infty$, dafür müssen wir zeigen, dass $\sqrt[n]{n!}$ unbeschränkt wächst. Für eine natürliche Zahl m suchen wir also ein N, so dass für alle $n > N$ gilt $\sqrt[n]{n!} > m$. Das ist äquivalent zu $n! > m^n$, bzw. $\prod_{k=1}^{n} \frac{k}{m} > 1$. Es ist leicht zu sehen, dass die Folge der Produkte $\prod_{k=1}^{n} \frac{k}{m}$ bestimmt gegen $+\infty$ divergiert. Damit hat die Exponentialreihe Konvergenzradius ∞. Die Behauptung folgt aus Satz 3.2.6. □

Bemerkung 5.1.4. *Der spezielle Wert* $\exp 1$ *erhält eine spezielle Bezeichnung und wird* Eulersche Zahl *genannt:*

$$\mathrm{e} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = \sum_{k=0}^{\infty} \frac{1}{k!}$$

Proposition 5.1.5 (Funktionalgleichung der Exponentialfunktion). *Für alle komplexen Zahlen $z, w \in \mathbb{C}$ gilt*

$$\exp(z + w) = \exp(z) \cdot \exp(w).$$

Beweis. Nach Proposition 5.1.3 ist die Exponentialreihe auf ganz $\mathbb{C}$ absolut konvergent. Mit dem binomischen Satz und der Formel für das Cauchy-Produkt aus Satz 2.6.24 haben wir

$$\begin{aligned}\exp(z+w) &= \sum_{m=0}^{\infty}\frac{(z+w)^m}{m!} = \sum_{m=0}^{\infty}\frac{1}{m!}\sum_{k=0}^{m}\binom{m}{k}z^k w^{m-k}\\ &= \sum_{m=0}^{\infty}\sum_{k=0}^{m}\frac{z^k w^{m-k}}{k!(m-k)!} = \left(\sum_{k=0}^{\infty}\frac{z^k}{k!}\right)\cdot\left(\sum_{l=0}^{\infty}\frac{w^l}{l!}\right)\\ &= \exp(z)\cdot\exp(w).\end{aligned}$$

□

Übungsaufgabe 5.1.6. *Wir betrachten eine Funktion $f(t)$, die einen exponentiellen Wachstums- oder Zerfallsprozess beschreibt, d.h. $f(t) = \lambda\exp(\alpha t)$. Zeigen Sie, dass die Wachstums- bzw. Zerfallsrate in einem Zeitintervall δ durch die folgende Formel bestimmt werden kann:*

$$\frac{f(t+\delta)}{f(t)} = \frac{f(t+2\delta)-f(t+\delta)}{f(t+\delta)-f(t)}$$

Bemerkung: *Auf diese Weise kann die effektive Reproduktionszahl für die Ausbreitung des Corona-Virus als Quotient der Neuinfektionszahlen in zwei aufeinanderfolgenden Zeitintervallen geschätzt werden. Die Länge des relevanten Zeitintervalls δ heißt in diesem Fall* Generationenzeit*. Schätzungen für die Generationenzeit schwanken zwischen 4 und 6 Tagen.*

Proposition 5.1.7. *Die Exponentialfunktion* $\exp\colon \mathbb{C}\to\mathbb{C}$ *ist stetig. Die Exponentialfunktion ist durch die folgenden beiden Eigenschaften eindeutig bestimmt:*

(1) $\exp(z+w) = \exp(z)\exp(w)$ *für alle* $z, w \in \mathbb{C}$

(2)
$$\lim_{z\to 0}\frac{\exp z - 1}{z} = 1$$

Beweis. Die Stetigkeit folgt aus der komplexen Version von Proposition 4.3.4 und Proposition 5.1.3. Für $z \neq 0$ haben wir

$$\frac{\exp z - 1}{z} = \sum_{k=1}^{\infty}\frac{z^{k-1}}{k!}.$$

Die Reihe auf der rechten Seite stellt eine auf ganz $\mathbb{C}$ stetige Funktion dar, deren Funktionswert $z = 0$ gleich 1 ist. Damit hat $\frac{\exp z-1}{z}$ eine stetige Fortsetzung auf $\mathbb{C}$, was die Grenzwertaussage liefert.

Sei nun f eine Funktion, die die beiden Eigenschaften erfüllt. Aus $f(z+w) = f(z)f(w)$ folgt

$$f(z) = \lim_{n\to\infty} f\left(\frac{z}{n}\right)^n.$$

Durch $f\left(\frac{z}{n}\right) = 1 + \frac{z_n}{n}$ definieren wir eine Folge (z_n), und aus $\lim_{z\to 0}\frac{f(z)-1}{z} = c$ folgt $\lim_{n\to\infty} z_n = cz$. Aus Proposition 5.1.1 folgt

$$f(z) = \lim_{n\to\infty}\left(1+\frac{z_n}{n}\right)^n = \lim_{n\to\infty}\left(1+\frac{cz}{n}\right)^n = \exp(cz).$$

□

Bemerkung 5.1.8. *Die Stetigkeit von* exp *kann auch direkter durch Restgliedabschätzung aus der Potenzreihendarstellung gezeigt werden.*

PROPOSITION 5.1.9. *Die reelle Exponentialfunktion* $\exp\colon \mathbb{R} \to \mathbb{R}$ *ist eine streng monoton wachsende Funktion. Für alle* $x \in \mathbb{R}$ *ist* $\exp(x) > 0$. *Die reelle Exponentialfunktion liefert eine Bijektion* $\exp\colon \mathbb{R} \to \mathbb{R}_+$.

BEWEIS. Aus der Definition folgt, dass für $x \in \mathbb{R}$ auch $\exp(x) \in \mathbb{R}$ ist. Aus der Funktionalgleichung folgt $\exp(x) = (\exp(\frac{x}{2}))^2$, damit ist $\exp(x) > 0$ für alle $x \in \mathbb{R}$, damit ist der Wertebereich in $\mathbb{R}_+$ enthalten.

Wir zeigen, dass der Wertebereich ganz $\mathbb{R}_+$ ist. Aus der Potenzreihendarstellung folgt $\exp(x) \geq 1 + x$, mit der Funktionalgleichung dann $\exp(-x) \leq \frac{1}{1+n}$. Daraus folgt $\lim_{x\to-\infty} \exp(x) = 0$ und $\lim_{x\to\infty} \exp(x) = \infty$. Nach dem Zwischenwertsatz 4.4.1 ist der Wertebereich $\mathbb{R}_+$.

Für die strenge Monotonie sei $x < y$. Dann ist $y - x > 0$ und damit $\exp(y-x) = 1 + (y-x) + \frac{(y-x)^2}{2} + \cdots > 1$. Mit der Funktionalgleichung sehen wir

$$\exp(y) = \exp(x) \cdot \exp(y-x) > \exp(x).$$

□

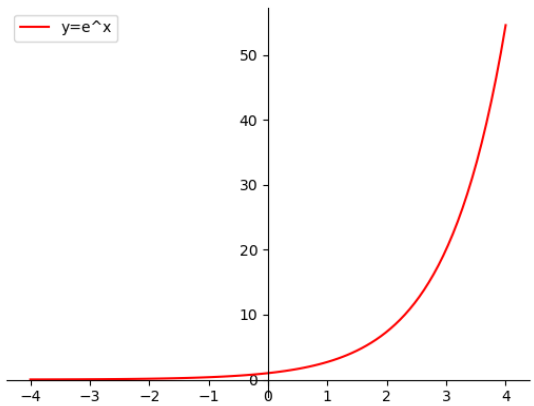


ABBILDUNG 15. Reelle Exponentialfunktion $x \mapsto \exp(x)$ im Intervall $[-4, 4]$

Die Exponentialfunktion wächst schneller als jedes Polynom:

ÜBUNGSAUFGABE 5.1.10. *Zeigen Sie:*

$$\lim_{x\to\infty} \frac{\mathrm{e}^x}{x^k} = \infty \qquad \lim_{x\to\infty} x^k \mathrm{e}^{-x} = 0$$

5.2. Logarithmus und allgemeine Potenz

Nach Proposition 5.1.9 ist die Exponentialfunktion streng monoton wachsend auf $\mathbb{R}$. Nach dem Umkehrsatz 4.5.2existiert eine stetige Umkehrfunktion, die *Logarithmusfunktion* $\ln\colon \mathbb{R}_+ \to \mathbb{R}$.

PROPOSITION 5.2.1. *Die Logarithmusfunktion erfüllt die Funktionalgleichung* $\ln(xy) = \ln x + \ln y$ *für alle* $x, y \in \mathbb{R}_+$.

BEWEIS. Wir setzen $a = \ln x$ und $b = \ln y$. Nach Definition des Logarithmus als Umkehrfunktion zur Exponentialfunktion ist $x = \exp a$ und $y = \exp b$ und die Funktionalgleichung der Exponentialfunktion liefert

$$\exp(a+b) = \exp(a)\exp(b) = xy.$$

Das bedeutet aber $\ln x + \ln y = a + b = \ln(xy)$. □

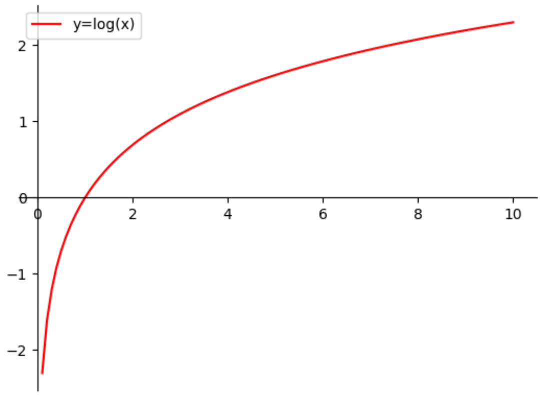


ABBILDUNG 16. Logarithmus $x \mapsto \ln(x)$ im Intervall $[0.1, 10]$

ÜBUNGSAUFGABE 5.2.2. *Zeigen Sie die folgenden Aussagen*

$$\lim_{x\to\infty} \ln x = \infty, \qquad \lim_{x\to 0^+} \ln x = -\infty$$

$$\lim_{x\to\infty} \frac{\ln x}{x^\alpha} = 0, \alpha \in \mathbb{Q}.$$

ÜBUNGSAUFGABE 5.2.3. *Zeigen Sie*

$$\lim_{x\to 0} \frac{\ln(1+x)}{x} = 1$$

ÜBUNGSAUFGABE 5.2.4. *Sei $a > 0$. Wir definieren rekursiv die Folge $(x_n)_n$ durch $x_0 = a$ und $x_{n+1} = \sqrt{x_n}$ und die Folge $(y_n)_n$ durch $y_n = 2^n(x_n - 1)$. Zeigen Sie, dass $\lim_{n\to\infty} y_n = \ln a$ gilt.*

Hinweis: *Benutzen Sie* $\lim_{x\to 0} \frac{\exp x - 1}{x} = 1$.

PROPOSITION 5.2.5. *Für alle rationalen Zahlen $\frac{p}{q} \in \mathbb{Q}$ gilt $\exp \frac{p}{q} = \sqrt[q]{e^p}$.*

BEWEIS. Für natürliche Zahlen $n \in \mathbb{N}$ gilt nach der Funktionalgleichung $\exp n = (\exp 1)^n = e^n$. Ebenfalls aus der Funktionalgleichung folgt $\exp(z)\exp(-z) = \exp 0 = 1$, damit folgt die Aussage für alle ganzen Zahlen. Für eine rationale Zahl $\frac{p}{q}$ gilt nach der Funktionalgleichung

$$e^p = \exp\left(q \cdot \frac{p}{q}\right) = \exp\left(\frac{p}{q}\right)^q$$

und daraus die Behauptung durch Ziehen der q-ten Wurzel. □

Für reelle Zahlen $a, x \in \mathbb{R}$ können wir mit Hilfe von Exponentialfunktion und Logarithmus die allgemeine Potenz a^x definieren:

$$a^x := \exp(x \ln a).$$

Mit dieser Definition von allgemeiner Potenz haben wir dann insbesondere $\exp(x) = e^x$. Für rationales x stimmt diese Definition nach Proposition 5.2.5 mit der üblichen Definition $a^{\frac{p}{q}} = \sqrt[q]{a^p}$ überein. Aus der Stetigkeit der Exponentialfunktion folgt, dass für jede Folge $(x_n)_n$ mit $\lim_{n\to\infty} x_n = x$ gilt $\lim_{n\to\infty} a^{x_n} = a^x$, s. auch Definition 2.4.2. Durch die Exponentialfunktion können wir auch eine Potenz a^z für $a \in \mathbb{R}$ und $z \in \mathbb{C}$ definieren.

ÜBUNGSAUFGABE 5.2.6. *Zeigen Sie*

$$\lim_{x\to 0^+} x^x = 1 \qquad \textit{und} \qquad \lim_{x\to 0^+} x\ln x = 0.$$

PROPOSITION 5.2.7. *Für die so definierte Potenzfunktion gelten die Potenzgesetze:*

(1) $a^x a^y = a^{x+y}$
(2) $(a^x)^y = a^{xy}$
(3) $a^x b^x = (ab)^x$
(4) $\left(\frac{1}{a}\right)^x = a^{-x}$

ÜBUNGSAUFGABE 5.2.8. *Für* $0 < x < 1$ *betrachten wir die geometrische Reihe* $\sum_{k=0}^n x^k$. *Wie groß muss* n *mindestens gewählt werden, damit sich die Partialsumme* $\sum_{k=0}^n x^k$ *von der Summe der Reihe* $\sum_{k=0}^\infty x^k$ *um weniger als einen vorgegebenen Fehler* ϵ *unterscheidet?*

Geben Sie den konkreten Zahlenwert für den Fall $x = 0.9$ *und* $\epsilon = 10^{-6}$ *an.*

Todo: Binomial und Logarithmusreihe

PROPOSITION 5.2.9. *Für reelle Zahlen* $a > 0$ *und* $b > 1$ *besitzt die Gleichung* $b^x = a$ *eine eindeutige Lösung, den Logarithmus* $\log_b a$ *von* a *zur Basis* b.

BEWEIS. Die Eindeutigkeit folgt aus der strengen Monotonie der Exponentialfunktion, s. Proposition 5.1.9. Wir beweisen die Existenz der Lösung. Aus $b > 1$ folgt $\lim_{n\to\infty} \left(\frac{1}{b}\right)^n = 0$. Da $a > 0$ ist, existiert ein N, so dass

$$\left(\frac{1}{b}\right)^N \leq \min\left(a, \frac{1}{a}\right).$$

Damit ist $b^{-N} \leq a \leq b^N$.

Nun definieren wir rekursiv eine Intervallschachtelung $I_n = [x_n, y_n]$ für $\log_b a$. Der Rekursionsanfang ist das Intervall $[x_1 = -N, y_1 = N]$. Im Rekursionsschritt halbieren wir $[x_k, y_k]$ durch den Mittelpunkt $m = \frac{x_k+y_k}{2}$ und setzen

$$[x_{k+1}, y_{k+1}] = \begin{cases} [x_k, m] & b^{x_k} \leq a \leq b^m \\ [m, y_k] & b^m < a \leq b^{y_k} \end{cases}$$

Diese Intervallschachtelung definiert eine eindeutige reelle Zahl x, so dass für die Intervallgrenzen x_k, y_k gilt $\lim_{k\to\infty} x_k = \lim_{k\to\infty} y_k = x$. Aus Monotonie und Stetigkeit der Potenzfunktion folgt $b^x \leq a \leq b^x$, was die Existenz der Lösung $\log_b a = x$ zeigt. □

5.3. Winkelfunktionen

DEFINITION 5.3.1. *Wir definieren die Winkelfunktionen für* $z \in \mathbb{C}$ *durch*

$$\cos(z) = \frac{\mathrm{e}^{\mathrm{i}z} + \mathrm{e}^{-\mathrm{i}z}}{2}, \qquad \sin(z) = \frac{\mathrm{e}^{\mathrm{i}z} - \mathrm{e}^{-\mathrm{i}z}}{2\mathrm{i}},$$

BEMERKUNG 5.3.2. *Für reelles* $t \in \mathbb{R}$ *liegt* $\mathrm{e}^{\mathrm{i}t}$ *wegen*

$$\|\mathrm{e}^{\mathrm{i}t}\|^2 = \mathrm{e}^{\mathrm{i}t} \cdot \overline{\mathrm{e}^{\mathrm{i}t}} = \mathrm{e}^{\mathrm{i}t} \cdot \mathrm{e}^{\overline{\mathrm{i}t}} = \mathrm{e}^{\mathrm{i}t} \cdot \mathrm{e}^{-\mathrm{i}t} = 1$$

auf dem Einheitskreis S^1. *Mit Hilfe der Übungsaufgabe 5.3.3 können wir auch zeigen* $\arg \mathrm{e}^{\mathrm{i}t} = t$. *Für reelles* $t \in \mathbb{R}$ *stimmt dann die obige Definition* $\cos(t) = \mathrm{Re}(\mathrm{e}^{\mathrm{i}t})$ *und* $\sin(t) = \mathrm{Im}(\mathrm{e}^{\mathrm{i}t})$ *mit der üblichen Definition der Winkelfunktionen am rechtwinkligen Dreieck überein. Es gilt die Eulersche Formel* $\mathrm{e}^{\mathrm{i}t} = \cos(t) + \mathrm{i}\sin(t)$.

ÜBUNGSAUFGABE 5.3.3. *Sei x eine reelle Zahl und $n \geq 1$ eine natürliche Zahl. Für $k = 0, 1, \ldots, n$ betrachten wir die Punkte $A_k^{(n)} = \exp\left(\frac{ikx}{n}\right)$. Sei*

$$L_n = \sum_{k=1}^{n} \left\| A_k^{(n)} - A_{k-1}^{(n)} \right\|$$

die Länge des Polygonzugs $A_0^{(n)}, A_1^{(n)}, \ldots, A_n^{(n)}$. Zeigen Sie

(a) $L_n = 2n \left|\sin \frac{x}{2n}\right|$.
(b) $\lim_{n\to\infty} 2n \sin \frac{x}{2n} = x$.

Die so definierten Winkelfunktionen sind wegen Proposition 5.1.7 stetig und es aus der Potenzreihendarstellung der Exponentialfunktion folgt

$$\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}, \qquad \sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

Die Potenzreihen konvergieren absolut und gleichmäßig auf $\mathbb{C}$.

ÜBUNGSAUFGABE 5.3.4. *Zeigen Sie* $\cos(-x) = \cos x$, $\sin(-x) = -\sin x$ *und* $\sin^2 x + \cos^2 x = 1$ *für* $x \in \mathbb{R}$.

PROPOSITION 5.3.5. *Für $z, w \in \mathbb{C}$ gilt*

$$\cos(z+w) = \cos z \cos w - \sin z \sin w, \qquad \sin(z+w) = \sin z \cos w + \cos z \sin w.$$

Als Spezialfälle haben wir $\cos(2z) = \cos^2 z - \sin^2 z$ *und* $\sin(2z) = 2 \sin z \cos z$.

Aus der Anwendung der Additionstheoreme auf $\frac{1}{2}(z+w) + \frac{1}{2}(z-w)$ *bzw.* $\frac{1}{2}(z+w) + \frac{1}{2}(w-z)$ *erhalten wir*

$$\cos z - \cos w = -2 \sin \frac{z+w}{2} \sin \frac{z-w}{2}, \qquad \sin z - \sin w = 2 \cos \frac{z+w}{2} \sin \frac{z-w}{2}.$$

BEWEIS. Folgt mit der Definition der Winkelfunktionen aus der Funktionalgleichung der Exponentialfunktion 5.1.5. □

Wir diskutieren kurz die Nullstellen und Periodizität der Winkelfunktionen (für reelle Argumente). Aus der alternierenden Potenzreihenentwicklung erhalten wir die folgenden Abschätzungen

$$1 - \frac{x^2}{2} < \cos x < 1 - \frac{x^2}{2} + \frac{x^4}{24}, \qquad x - \frac{x^3}{6} < \sin x < x.$$

Daraus folgt $\sin x > 0$ für $x \in (0, 2]$. Aus dem Additionstheorem für $\cos z - \cos w$ folgt, dass $\cos x$ auf $[0, 2]$ streng monoton fallend ist. Es gilt $\cos 0 = 1$, und aus der obigen Abschätzung ist $\cos 2 < -\frac{1}{3}$. Nach dem Zwischenwertsatz 4.4.1 hat $\cos x$ in $[0, 2]$ eine Nullstelle. Wegen der strengen Monotonie ist diese Nullstelle eindeutig und wird mit $\frac{\pi}{2}$ bezeichnet. Aus dem trigonometrischen Pythagoras folgt $\sin^2 \frac{\pi}{2} = 1$, und da $\sin x$ auf $(0, 2]$ positiv ist, folgt $\sin \frac{\pi}{2} = 1$.

Daraus folgt $\exp\left(\frac{\mathrm{i}\pi}{2}\right) = \mathrm{i}$ und durch Potenzieren

$$\exp(\pi \mathrm{i}) = -1, \qquad \exp(2\pi \mathrm{i}) = 1.$$

Aus der Funktionalgleichung für die Exponentialfunktion folgt, dass die Exponentialfunktion die komplexe Periode $2\pi\mathrm{i}$ besitzt, d.h. für alle $z \in \mathbb{C}$ gilt $\exp(z + 2\pi\mathrm{i}) = \exp(z)$. Daraus folgt, dass die Winkelfunktionen sin und cos die reelle Periode 2π haben. Die Menge der Nullstellen der Sinusfunktion ist $\{k\pi \mid k \in \mathbb{Z}\}$, die Menge der Nullstellen der Cosinusfunktion ist $\left\{\frac{(2k+1)\pi}{2} \mid k \in \mathbb{Z}\right\}$.

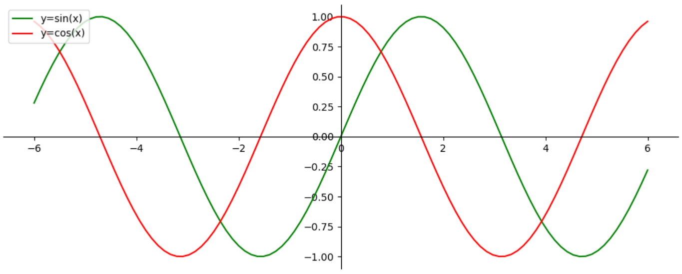


ABBILDUNG 17. Winkelfunktionen im Intervall $[-6, 6]$

DEFINITION 5.3.6. *Wir definieren die Tangens-Funktion*

$$\tan\colon \mathbb{R} \setminus \{z \in \mathbb{R} \mid \cos(z) = 0\} \to \mathbb{R}\colon x \mapsto \tan x = \frac{\sin x}{\cos x}.$$

und die Cotangens-Funktion

$$\cot\colon \mathbb{R} \setminus \{z \in \mathbb{R} \mid \sin(z) = 0\} \to \mathbb{R}\colon x \mapsto \cot x = \frac{\cos x}{\sin x}.$$

ÜBUNGSAUFGABE 5.3.7. *Zeigen Sie, dass der Tangens das folgende Additionstheorem erfüllt: für $x, y \in \mathbb{R}$, so dass $\tan x$, $\tan y$ und $\tan(x+y)$ definiert sind, gilt*

$$\tan(x+y) = \frac{\tan x + \tan y}{1 - \tan x \tan y}.$$

ÜBUNGSAUFGABE 5.3.8. *Die auf $\mathbb{C} \setminus \mathbb{Z}$ definierte Funktion*

$$g(z) = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2}$$

ist stetig und periodisch mit Periode 1: $g(z+1) = g(z)$.

Die Sinusfunktion ist auf dem abgeschlossenen Intervall $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ streng monoton wachsend, nach dem Zwischenwertsatz ist der Wertebereich $[-1, 1]$. Folglich existiert eine stetige Umkehrfunktion

$$\arcsin\colon [-1, 1] \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Die Cosinusfunktion ist auf dem abgeschlossenen Intervall $[0, \pi]$ streng monoton fallend, nach dem Zwischenwertsatz ist der Wertebereich $[-1, 1]$. Damit existiert eine stetige Umkehrfunktion

$$\arccos\colon [-1, 1] \to [0, \pi].$$

Die Tangensfunktion ist auf dem offenen Intervall $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ stetig und monoton wachsend, mit Wertebereich $\mathbb{R}$. Damit existiert eine stetige Umkehrfunktion

$$\arctan\colon \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

Todo: Arkustangens-Reihe und Leibniz-Reihe für $\frac{\pi}{4}$

Zum Schluss diskutieren wir noch die Polarkoordinaten komplexer Zahlen und ihre Darstellung mit Hilfe der komplexen Exponentialfunktion.

PROPOSITION 5.3.9. *Jede komplexe Zahl $z \neq 0$ besitzt eine Darstellung $z = r\exp(i\phi)$ mit $r = \|z\|$ und $\phi \in \mathbb{R}$. Dabei ist ϕ bis auf die Addition eines ganzzahligen Vielfachen von 2π eindeutig bestimmt.*

Beweis. Ersetzen wir z durch $\frac{z}{\|z\|}$ reicht es, die Aussage für komplexe Zahlen mit Betrag 1 zu zeigen. Sei also $z = a + b\mathrm{i}$ mit $a^2 + b^2 = 1$.

Wir betrachten zuerst den Fall $b \geq 0$, der Fall $b \leq 0$ folgt daraus durch komplexe Konjugation. Wir setzen $\phi = \arccos a$. Dann ist $\phi \in [0, \pi]$ und wir haben $a = \cos\phi$ und $\sin\phi \geq 0$. Aus $a^2 + b^2 = 1$ und $b \geq 0$ folgt $b = \sin\phi$. Wir erhalten wie gewünscht

$$a + b\mathrm{i} = \cos\phi + \mathrm{i}\sin\phi = \exp(\mathrm{i}\phi).$$

Sei nun $z = \|z\| \exp(\mathrm{i}\phi) = \|z\| \exp(\mathrm{i}\psi)$. Dann ist $\exp(\mathrm{i}(\phi - \psi)) = 1$, und daraus folgt $\phi - \psi = 2\pi n$ für ein $n \in \mathbb{Z}$. □

Bemerkung 5.3.10. *Die Winkelkoordinate ϕ heißt auch* Argument *der komplexen Zahl z und wird mit* $\arg z$ *bezeichnet.*

Übungsaufgabe 5.3.11. *Zeigen Sie, dass für eine komplexe Zahl $z = r\exp(\mathrm{i}\phi)$ die n-ten Wurzeln durch die folgende Formel gegeben sind:*

$$z_k = \sqrt[n]{r}\exp\left(\mathrm{i}\frac{\phi + 2\pi k}{n}\right), \qquad k = 0, 1, \ldots, n-1.$$

Übungsaufgabe 5.3.12. *Zeigen Sie*

$$\sum_{k=1}^{n}\left(\cos\left(\frac{2\pi k}{n}\right) + \mathrm{i}\sin\left(\frac{2\pi k}{n}\right)\right) = 0$$

und interpretieren Sie die Aussage geometrisch (Zeichnung im Fall $n = 3$).

Hinweis: *Beweisen Sie durch vollständige Induktion, dass für alle natürlichen Zahlen $k \geq 1$ und alle reellen Zahlen ϕ gilt*

$$\cos(k\phi) + \mathrm{i}\sin(k\phi) = (\cos\phi + \mathrm{i}\sin\phi)^k.$$

Wenden Sie dann die Formel für die geometrische Summe an.

Todo: hyperbolischer Sinus und Cosinus, Kettenlinie

KAPITEL 6

Differentialrechnung in einer Variablen

6.1. Ableitung und Differenzierbarkeit: Definition und Beispiele

Die Ableitung einer Funktion $f\colon D \to \mathbb{R}$ beschreibt (wenn sie existiert) die Änderungsrate von f im Punkt p.

DEFINITION 6.1.1. *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $f\colon I \to \mathbb{R}$ eine reellwertige Funktion und $x_0 \in I$. Die Funktion f ist* differenzierbar im Punkt x_0*, wenn der Grenzwert*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert. Der Grenzwert wird dann mit $f'(x_0)$ bezeichnet.

Die Funktion f heißt differenzierbar *auf I, wenn sie in allen Punkten $x_0 \in I$ differenzierbar ist. In diesem Fall haben wir eine Funktion $f'\colon I \to \mathbb{R}\colon x \mapsto f'(x)$. Die Funktion f heißt* stetig differenzierbar*, wenn die Ableitung f' auf I existiert und stetig ist.*

BEMERKUNG 6.1.2. *Der Differenzenquotient in der Definition wird auch als $\frac{\Delta f}{\Delta x}$ geschrieben, die Ableitung wird dann alternativ als $\frac{\mathrm{d}f}{\mathrm{d}x}$ geschrieben. In der Physik ist auch die Notation $\dot{f} = \frac{\mathrm{d}f}{\mathrm{d}t}$ für Ableitungen nach der Zeit gebräuchlich.*

BEISPIEL 6.1.3. *(1) Die Ableitung von $f(x) = x^n$ für natürliche Zahlen n ist nx^{n-1}. Mit einer Variation der geometrischen Summe und der Stetigkeit von Polynomen haben wir*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{x^n - x_0^n}{x - x_0} = \lim_{x \to x_0} \left(\sum_{i=0}^{n-1} x^i x_0^{n-1-i} \right) = nx^{n-1}.$$

(2) Die Ableitung von $f(x) = \exp(\lambda x)$ ist $f'(x) = \lambda \exp(\lambda x)$. Das folgt aus den Eigenschaften der Exponentialfunktion, s. Proposition 5.1.7:

$$\begin{aligned} \lim_{x \to x_0} \frac{\exp(\lambda(x + x_0 - x_0)) - \exp(\lambda x_0)}{x - x_0} &= \exp(\lambda x_0) \lim_{x \to x_0} \frac{\exp(\lambda(x - x_0)) - 1}{x - x_0} \\ &= \lambda \exp(\lambda x_0). \end{aligned}$$

(3) Die Ableitung des Logarithmus $f(x) = \ln(x)$ ist die Funktion $f'(x) = \frac{1}{x}$, durch Benutzung der Funktionalgleichung Proposition 5.2.1 und Übungsaufgabe 5.2.3:

$$\begin{aligned} \lim_{x \to x_0} \frac{\ln(x - x_0 + x_0) - \ln(x_0)}{x - x_0} &= \frac{1}{x_0} \lim_{x \to x_0} \frac{\ln\left(x_0 \left(1 + \frac{x - x_0}{x_0}\right)\right) - \ln(x_0)}{\frac{x - x_0}{x_0}} \\ &= \frac{1}{x_0} \lim_{x \to x_0} \frac{\ln\left(1 + \frac{x - x_0}{x_0}\right)}{\frac{x - x_0}{x_0}} \\ &= \frac{1}{x_0} \end{aligned}$$

□

ÜBUNGSAUFGABE 6.1.4. *Zeigen Sie* $\frac{\mathrm{d}\sin x}{\mathrm{d}x} = \cos x$, *unter Benutzung der Additionstheoreme und der Berechnung von* $\lim_{h\to 0} \frac{\sin h}{h}$ *aus Beispiel 4.2.5 (2).*

BEMERKUNG 6.1.5. *Die Ableitung kann geometrisch interpretiert werden. Der Differenzenquotient* $\frac{f(x)-f(x_0)}{x-x_0}$ *ist der Anstieg einer Sekante des Funktionsgraphen durch die Punkte* $(x, f(x))$ *und* $(x, f(x_0))$. *Die Ableitung als Grenzwert* $x \to x_0$ *des Differenzenquotienten ist der Anstieg der Tangente im Punkt* x_0. *Die Gleichung für die Tangent an den Graphen von* f *im Punkt* x_0 *ist*

$$y = f'(x_0)(x - x_0) + f(x_0).$$

Die lineare Funktion $x \mapsto f'(x_0)(x - x_0) + f(x_0)$ *ist die beste lineare Annäherung an die Funktion* f.

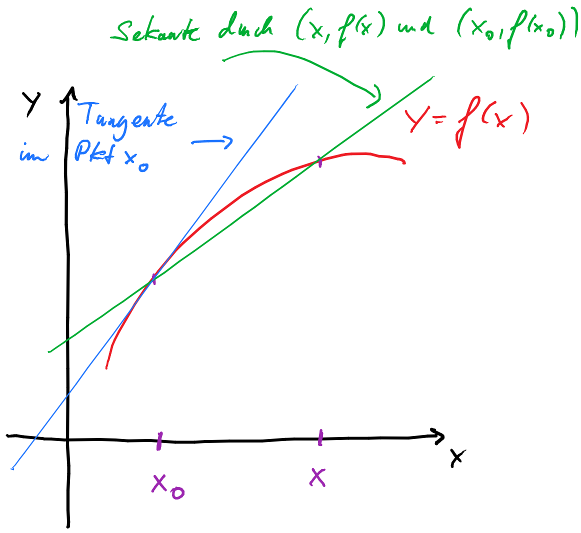


ABBILDUNG 18. Geometrische Interpretation der Ableitung als Anstieg der Tangente bzw. lineare Annäherung.

ÜBUNGSAUFGABE 6.1.6. *Ermitteln Sie die Gleichungen der Tangenten an den Graphen der Funktion* $f(x) = x^3 + x^2 - x + 1$ *in den Punkten* $x = -1$, $x = 0$ *und* $x = 1$.

PROPOSITION 6.1.7. *Eine differenzierbare Funktion ist stetig.*

BEWEIS. Wenn f an der Stelle x_0 differenzierbar ist, haben wir nach den Rechenregeln für Produkte konvergenter Folgen

$$0 = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} \cdot \lim_{x\to x_0} (x - x_0) = \lim_{x\to x_0} (f(x) - f(x_0)).$$

□

BEMERKUNG 6.1.8. *Die Umkehrung gilt nicht. Standardbeispiel ist die Betragsfunktion* $f(x) = |x|$, *die stetig aber im Punkt* 0 *nicht differenzierbar ist.*

DEFINITION 6.1.9. *Höhere Ableitungen werden (so sie denn existieren) rekursiv definiert:*

$$f^{(0)}(x) = f(x), \qquad f^{(n)}(x) = \frac{\mathrm{d}^n f(x)}{\mathrm{d}x^n} := \frac{\mathrm{d}f^{(n-1)}(x)}{\mathrm{d}x}.$$

6.2. Ableitungsregeln

PROPOSITION 6.2.1. *Seien $f, g\colon I \to \mathbb{R}$ differenzierbare Funktionen. Dann gelten die folgenden Rechenregeln:*

(1) $(f(x) \pm g(x))' = f'(x) \pm g'(x)$
(2) $(f(x) \cdot g(x))' = f'(x)g(x) + f(x)g'(x)$
(3) Wenn $g(x) \neq 0$ ist, dann gilt

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}$$

BEWEIS. (1) Wir schreiben den Differenzenquotienten geeignet um:

$$\lim_{x\to x_0} \frac{f(x) \pm g(x) - f(x_0) \mp g(x_0)}{x - x_0} = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} \pm \lim_{x\to x_0} \frac{g(x) - g(x_0)}{x - x_0}$$

Die Behauptung folgt dann aus den Grenzwertrechenregeln.

(2) Wieder geeignetes Umschreiben des Differenzenquotienten

$$\begin{aligned}
\frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} &= \frac{f(x)g(x) + f(x_0)g(x) - f(x_0)g(x) - f(x_0)g(x_0)}{x - x_0} \\
&= f(x_0)\frac{g(x) - g(x_0)}{x - x_0} + g(x)\frac{f(x) - f(x_0)}{x - x_0} \\
\lim_{x\to x_0} f(x_0)\frac{g(x) - g(x_0)}{x - x_0} &= f(x_0)g'(x_0) \\
\lim_{x\to x_0} g(x)\frac{f(x) - f(x_0)}{x - x_0} &= g(x_0)f'(x_0)
\end{aligned}$$

(3) analog zu (2) mit

$$\frac{\frac{f(x)}{g(x)} - \frac{f(x_0)}{g(x_0)}}{x - x_0} = \frac{1}{g(x)g(x_0)}\left(\frac{f(x) - f(x_0)}{x - x_0}g(x_0) - \frac{g(x) - g(x_0)}{x - x_0}f(x_0)\right)$$

□

BEISPIEL 6.2.2. *(1) Für ein Polynom $p(x) = \sum_{i=0}^n a_i x^i$ haben wir*

$$p'(x) = \sum_{i=1}^n a_i i x^{i-1}$$

(2) Nach der Quotientenregel gilt

$$\left(\frac{1}{x^n}\right)' = -\frac{n}{x^{n+1}}$$

(3) Rationale Funktionen sind auf ihrem gesamten Definitionsbereich differenzierbar, und die Ableitung einer rationalen Funktion ist wieder eine rationale Funktion.

(4) Nach Definition der Winkelfunktionen ist

$$(\cos x)' = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\mathrm{e}^{\mathrm{i}x} + \mathrm{e}^{-\mathrm{i}x}}{2} = \frac{\mathrm{i}(\mathrm{e}^{\mathrm{i}x} - \mathrm{e}^{-\mathrm{i}x})}{2} = -\sin x.$$

Analog folgt $(\sin x)' = \cos x$. Aus der Quotientenregel folgt

$$(\tan x)' = \frac{(\sin x)' \cdot \cos x - \sin x \cdot (\cos x)'}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}$$

□

Proposition 6.2.3 (Kettenregel). *Seien $g\colon I \to \mathbb{R}$ und $f\colon J \to \mathbb{R}$ zwei differenzierbare Funktionen mit $g(I) \subseteq J$. Dann ist die Komposition $f \circ g(x) = f(g(x))$ differenzierbar und es gilt*

$$(f(g(x)))' = f'(g(x)) \cdot g'(x).$$

Beweis. Wir betrachten die Funktion

$$D\colon J \to \mathbb{R}\colon y \mapsto D(y, y_0) = \begin{cases} \frac{f(y)-f(y_0)}{y-y_0} & y \neq y_0 \\ f'(y_0) & y = y_0 \end{cases}$$

Diese Funktion ist stetig, da f differenzierbar ist und es gilt $f(y) - f(y_0) = D(y, y_0)(y - y_0)$. Dann haben wir

$$\begin{aligned} \lim_{x\to x_0} \frac{f(g(x)) - f(g(x_0))}{x - x_0} &= \lim_{x\to x_0} \frac{D(g(x), g(x_0))\,(g(x) - g(x_0))}{x - x_0} \\ &= \lim_{x\to x_0} D(g(x), g(x_0)) \lim_{x\to x_0} \frac{g(x) - g(x_0)}{x - x_0} \\ &= f'(g(x_0))g'(x_0). \end{aligned}$$

□

Beispiel 6.2.4.
(1) Für die allgemeine Potenzfunktion $f\colon \mathbb{R}_+ \to \mathbb{R}\colon x \mapsto x^a$, $a \in \mathbb{R}$ gilt

$$\frac{\mathrm{d}x^a}{\mathrm{d}x} = \frac{\mathrm{d}\exp(a \ln x)}{\mathrm{d}x} = \exp(a \ln x) \cdot \frac{a}{x} = ax^{a-1}.$$

(2) Allgemeiner ist für eine positive differenzierbare Funktion $f\colon I \to \mathbb{R}$ und $a \in \mathbb{R}$ auch f^a differenzierbar und es gilt $(f^a)' = af^{a-1} \cdot f'$.
(3) $\exp(f(x))' = \exp(f(x)) \cdot f'(x)$.

□

Proposition 6.2.5 (Umkehrfunktion). *Sei $f\colon I \to \mathbb{R}$ eine stetige, streng monotone Funktion und sei $f^{-1}\colon f(I) \to \mathbb{R}$ die Umkehrfunktion. Wenn f im Punkt $x \in I$ differenzierbar ist und $f'(x) \neq 0$, dann ist f^{-1} im Punkt $y = f(x) \in f(I)$ differenzierbar und es gilt*

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}.$$

Beweis. Nach dem Umkehrsatz 4.5.2 ist f^{-1} stetig. Sei (y_n) eine Folge in $f(I)$ mit $\lim_{n\to\infty} y_n = f(x)$. Dann gilt $\lim_{n\to\infty} f^{-1}(y_n) = f^{-1}(f(x)) = x$ und wir haben

$$\begin{aligned} \lim_{n\to\infty} \frac{f^{-1}(y_n) - f^{-1}(f(x))}{y_n - f(x)} &= \lim_{n\to\infty} \frac{f^{-1}(y_n) - x}{f(f^{-1}(y_n)) - f(x)} \\ &= \lim_{n\to\infty} \frac{1}{\frac{f(f^{-1}(y_n)) - f(x)}{f^{-1}(y_n) - x}} \\ &= \frac{1}{f'(x)} \end{aligned}$$

Damit folgt die Behauptung. □

Bemerkung 6.2.6. *Die Formel folgt aus der Kettenregel für die Komposition $f \circ f^{-1} = \mathrm{id}$. Leiten wir die Gleichung auf beiden Seiten ab, erhalten wir*

$$f'(f^{-1}(x)) \cdot (f^{-1})'(x) = 1.$$

Umstellen liefert die Behauptung. Die Differenzierbarkeit der Umkehrfunktion ist allerdings Voraussetzung für die Anwendung der Kettenregel.

BEISPIEL 6.2.7. *(1) Die Wurzelfunktion $x \mapsto \sqrt[n]{x}$ ist die Umkehrfunktion der Potenzfunktion $x \mapsto x^n$. Für gerade n ist der Definitionsbereich der Umkehrfunktion $\mathbb{R}_{\geq 0}$. Mit der Notation $x^{\frac{1}{n}} = \sqrt[n]{x}$ haben wir dann*

$$(x^\alpha)' = \alpha x^{\alpha-1}.$$

(2) Die Formel für die Ableitung der Umkehrfunktion kann auch benutzt werden, um die Ableitung des natürlichen Logarithmus auszurechnen:

$$(\ln x)' = \frac{1}{\exp'(\ln x)} = \frac{1}{\exp(\ln x)} = \frac{1}{x}.$$

(3) Auf dem Intervall $[-\frac{\pi}{2}, \frac{\pi}{2}]$ ist $\sin(x)$ umkehrbar, die Umkehrfunktion ist $\arcsin\colon [-1,1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$. Nach der Rechenregel für Umkehrfunktion ist

$$\arcsin'(x) = \frac{1}{\cos(\arcsin(x))} = \frac{1}{\sqrt{1-\sin^2(\arcsin(x))}} = \frac{1}{\sqrt{1-x^2}}.$$

Analog ist die Funktion $\cos(x)$ auf $[0,\pi]$ umkehrbar mit Umkehrfunktion $\arccos\colon [-1,1] \to [0,\pi]$ und es gilt

$$\arccos'(x) = -\frac{1}{\sin(\arccos(x))} = -\frac{1}{\sqrt{1-\cos^2(\arccos(x))}} = -\frac{1}{\sqrt{1-x^2}}.$$

□

ÜBUNGSAUFGABE 6.2.8. *Leiten Sie die Formel*

$$\arctan'(x) = \frac{1}{1+x^2}$$

nur unter Verwendung der Kettenregel für die Ableitung und den bekannten Ableitungen für $\sin$ und $\cos$, sowie der Relation $\sin^2(x) + \cos^2(x) = 1$ her.

ÜBUNGSAUFGABE 6.2.9. *Bestimmen Sie die Ableitungen der folgenden Funktionen*

(1) $f(x) = \frac{x^2-1}{x^2+1}$
(2) $f(x) = \mathrm{e}^x \cdot \ln(x^5)$
(3) $f(x) = \ln\left(\sqrt{\frac{x+1}{x-1}}\right)$
(4) $f_\alpha(x) = \ln(\alpha \exp(x))$
(5) $f(x) = \exp\left(\frac{1}{1+x^2}\right)$
(6) $f(x) = \frac{x^5+10x^4+3x+7}{x^7+8x^3+9}$
(7) $f_n(x) = \frac{1}{\sqrt[2n]{x^{2n}+1}}$
(8) $f_n(x) = \underbrace{\ln(\ln(\cdots\ln(}_{n\text{-mal}}x)\cdots)),\ n \in \mathbb{N}$
(9) $f(x) = \ln\left(\frac{1}{sinx} - \frac{1}{\tan x}\right)$

ÜBUNGSAUFGABE 6.2.10. *Gegeben sei eine reelle Zahl $a > 0$. Berechnen Sie die Ableitungen der folgenden Funktionen $f_i\colon \mathbb{R}_+ \to \mathbb{R}$:*

$$f_1(x) = x^{(x^x)}, \qquad f_2(x) = (x^x)^x, \qquad f_3(x) = x^{(x^a)}, \qquad f_4(x) = x^{(a^x)}, \qquad f_5(x) = a^{(x^x)}.$$

ÜBUNGSAUFGABE 6.2.11. *Gegeben sei die Funktion*

$$\alpha\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto \begin{cases} 0 & x \leq 0 \\ \mathrm{e}^{-\frac{1}{x}} & x > 0 \end{cases}$$

Zeigen Sie durch vollständige Induktion, dass es Polynome $P_n(T)$ gibt, so dass

$$\alpha^{(n)}(x) = \begin{cases} P_n(\frac{1}{x})\,\mathrm{e}^{-\frac{1}{x}} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Insbesondere ist die Funktion $\alpha(x)$ beliebig of stetig differenzierbar.

Hinweis: *Benutzen Sie für auftretende Grenzwertbetrachtungen, dass die Exponentialfunktion schneller wächst als jedes Polynom.*

ÜBUNGSAUFGABE 6.2.12. *Seien $f, g\colon I \to \mathbb{R}$ zwei stetig differenzierbare Funktionen. Berechnen Sie die Ableitung der Funktion $f(x)^{g(x)}$. Welche Voraussetzungen müssen dafür an f bzw. g gestellt werden?*

ÜBUNGSAUFGABE 6.2.13. *Der Radius einer Kugel wird mit einer Genauigkeit von 1% bestimmt. Schätzen Sie mit Hilfe der Differentialrechnung den relativen Fehler bei der Berechnung des Kugelvolumens ab.*

ÜBUNGSAUFGABE 6.2.14. *Die hyperbolische Kosinusfunktion ist gegeben durch $\cosh(x) = \frac{\exp(x)+\exp(-x)}{2}$. Zeigen Sie, dass die Funktion $y(x) = a \cdot \cosh\left(\frac{x-x_0}{a}\right) + y_0$ die folgende Gleichung erfüllt:*

$$(y - y_0)y'' - (y')^2 = 1.$$

ÜBUNGSAUFGABE 6.2.15. *Ein Massepunkt schwingt nach $x(t) = A \sin \omega t$ um seine Ruhelage. Zeigen Sie, dass die Bewegung die Gleichung $x''(t) + \omega^2 x(t) = 0$ erfüllt.*

6.3. Mittelwertsatz, Monotonie und Konvexität

6.3.1. Lokale Extrema und Mittelwertsatz.

DEFINITION 6.3.1 (globale bzw. lokale Extremstellen). *Sei $D \subseteq \mathbb{R}$ und $f\colon D \to \mathbb{R}$ eine reellwertige Funktion.*

- *Die Funktion f hat in $a \in D$ ein* globales Maximum *(bzw.* globales Minimum*), wenn für alle $x \in D$ gilt $f(x) \leq f(a)$ (bzw.$f(x) \geq f(a)$).*
- *Die Funktion f hat in $b \in D$ ein* lokales Maximum *(bzw.* lokales Minimum*), wenn es eine Umgebung $(b - \epsilon, b + \epsilon)$ on b gibt, so dass für alle $x \in D \cap (b - \epsilon, b + \epsilon)$ gilt $f(x) \leq f(a)$ (bzw.$f(x) \geq f(a)$).*

PROPOSITION 6.3.2. *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und $f\colon I \to \mathbb{R}$ differenzierbar. Wenn $x_0 \in I$ ein lokales Extremum ist, dann gilt $f'(x_0) = 0$.*

BEWEIS. Sei x_0 ein lokales Maximum, d.h. $f(x) - f(x_0) \leq 0$ für alle $x \in (x_0 - \epsilon, x_0 + \epsilon)$. Dann ist

$$\frac{f(x) - f(x_0)}{x - x_0} \begin{cases} \leq 0 & x > x_0 \\ \geq 0 & x < x_0. \end{cases}$$

Damit gilt

$$\lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \leq 0, \qquad \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \geq 0$$

Da der Grenzwert existiert, stimmen rechtsseitiger und linksseitiger Grenzwert überein und wir erhalten für die Ableitung

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = 0.$$

□

BEMERKUNG 6.3.3. *Für eine Funktion f kommen als mögliche Extremstellen also in Frage:*

- *Randpunkte des Definitionsbereichs,*
- *Punkte, an denen f nicht differenzierbar ist,*
- *Punkte, an denen $f'(x_0) = 0$ ist. Solche Punkte heißen* stationäre/kritische Punkte.

Satz 6.3.4 (Mittelwertsatz). *Die Funktion $f\colon [a,b] \to \mathbb{R}$ sei stetig auf $[a,b]$ und differenzierbar auf (a,b). Dann existiert ein $x_0 \in (a,b)$ mit*

$$f'(x_0) = \frac{f(b) - f(a)}{b-a}.$$

Beweis. Wir betrachten die Funktion

$$F(x) := f(x) - (x-b)\frac{f(b)-f(a)}{b-a}.$$

Diese Funktion ist nach Voraussetzung stetig auf $[a,b]$ und differenzierbar auf (a,b) und es gilt $F(a) = F(b) = f(b)$. Nach Satz 4.4.7 existiert ein Minimum und ein Maximum für $F(x)$. Wenn die Funktion $F(x)$ konstant ist, ist $f(x)$ linear, und die Behauptung gilt für alle $x \in (a,b)$. Wenn $F(x)$ nicht konstant ist, liegt ein Extremum x_0 von F in (a,b). Damit gilt

$$F'(x_0) = 0 = f'(x_0) - \frac{f(b)-f(a)}{b-a}.$$

□

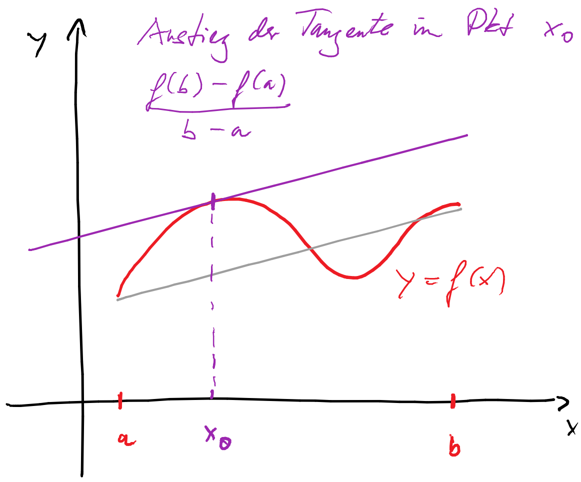


Abbildung 19. Mittelwertsatz: Es existiert ein Punkt x_0, in dem der Anstieg der Tangente an den Funktionsgraphen $y = f(x)$ gleich $\frac{f(b)-f(a)}{b-a}$ ist.

Korollar 6.3.5. *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und $f\colon I \to \mathbb{R}$ differenzierbar. Dann gelten die folgenden Aussagen:*

(1) Wenn $f'(x) \geq 0$ (bzw. > 0) auf I ist, dann ist f (streng) monoton wachsend.
(2) Wenn $f'(x) \leq 0$ (bzw. < 0) auf I ist, dann ist f (streng) monoton fallend.
(3) Wenn $f'(x) = 0$ auf I ist, dann ist f konstant.

Beweis. (1) Für $a, b \in I$ folgt aus dem Mittelwertsatz, dass ein $x_0 \in [a,b]$ existiert, so dass

$$\frac{f(b)-f(a)}{b-a} = f'(x_0) > 0$$

gilt. Daraus folgt $f(b) > f(a)$.

Die anderen Fälle werden analog bewiesen. □

Übungsaufgabe 6.3.6. *(a) Seien $f, g\colon [a,b] \to \mathbb{R}$ stetige auf (a,b) differenzierbare Funktionen mit $f(a) \geq g(a)$ und $f' \geq g'$ auf (a,b). Zeigen Sie, dass $f \geq g$ auf $[a,b]$ gilt.*

(b) Zeigen Sie als Anwendung

$$1 - \frac{1}{x} \leq \ln x \leq x - 1$$

für alle $x > 0$.

Korollar 6.3.7. *Sei $(a,b) \subseteq \mathbb{R}$ ein offenes Intervall, $f\colon (a,b) \to \mathbb{R}$ eine differenzierbare Funktion und $x_0 \in (a,b)$ ein kritischer Punkt. Dann ist x_0 ein lokales Maximum (bzw. Minimum), wenn ein $\epsilon > 0$ existiert, so dass $f'(x) > 0$ (bzw. $f'(x) < 0$) für alle $x \in (x_0 - \epsilon, x_0)$ gilt und $f'(x) < 0$ (bzw. $f'(x) > 0$) für alle $x \in (x_0, x_0 + \epsilon)$ gilt.*

Wenn f auf (a,b) zweimal stetig differenzierbar ist, kann die Bedindung durch $f''(x_0) < 0$ (bzw. $f''(x_0) > 0$) ersetzt werden.

Beweis. Folgt direkt aus den Monotoniekriterien. Für die Umformulierung mit der zweiten Ableitung sehen wir, dass aus $f''(x_0) < 0$ folgt, dass ein ϵ existiert, so dass $f''(x) < 0$ für alle x in einer ϵ-Umgebung von x_0 gilt, s. Übungsaufgabe 4.1.8. Nach dem Monotoniekriterium ist dann $f'(x)$ in dieser ϵ-Umgebung streng monoton fallend. Daraus folgt $f'(x) > 0$ für $x \in (x_0 - \epsilon, x_0)$ und $f'(x) < 0$ für $x \in (x_0, x_0 + \epsilon)$. □

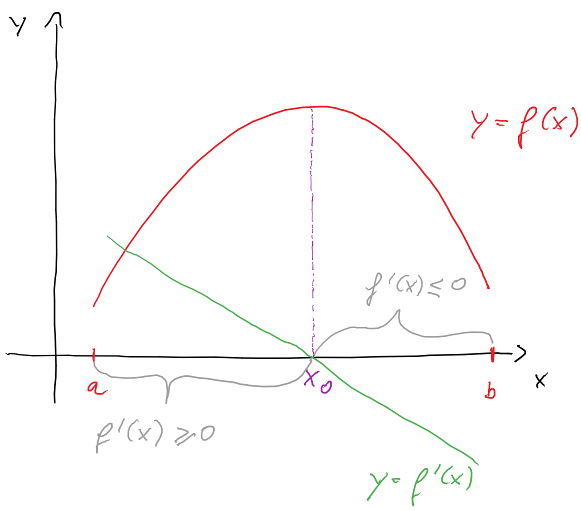


Abbildung 20. Kriterium für Existenz von lokalem Maximum

Übungsaufgabe 6.3.8. *Zeigen Sie, dass eine stetig differenzierbare Funktion $f\colon [a,b] \to \mathbb{R}$ auf einem abgeschlossenen Intervall Lipschitz-stetig ist, d.h. es existiert ein L, so dass für alle $x, y \in [a,b]$ gilt $|f(x) - f(y)| \leq L|x - y|$.*

Proposition 6.3.9 (Verallgemeinerter Mittelwertsatz). *Seien $f, g\colon [a,b] \to \mathbb{R}$ stetig und im Intervall (a,b) differenzierbar, und es gelte $g'(x) \neq 0$ für alle $x \in (a,b)$. Dann ist $g(b) \neq g(a)$ und es gibt ein $x_0 \in (a,b)$, so dass gilt*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x_0)}{g'(x_0)}.$$

Beweis. Wenn $g(b) = g(a)$ gilt, folgt aus dem Mittelwertsatz 6.3.4, dass es ein $x_0 \in (a, b)$ gibt, an dem $g'(x_0) = 0$ gilt. Aus der Voraussetzung $g'(x) \neq 0$ für alle $x \in (a, b)$ folgt also $g(b) \neq g(a)$. Wir betrachten die Funktion

$$F(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)} \left(g(x) - g(a)\right).$$

Dann ist $F(b) = F(a)$, und nach dem Mittelwertsatz gibt es ein $x_0 \in (a, b)$ mit $F'(x_0) = 0$. Daraus folgt die Behauptung. □

Beispiel 6.3.10. *Planck-Strahlungsgesetz* □

Übungsaufgabe 6.3.11. *Welches Rechteck mit gegebenem Umfang U hat die größte Fläche?*

Übungsaufgabe 6.3.12. *Für welche Kantenlänge x und y hat ein Rechteck mit vorgegebener Diagonallänge die maximale Fläche?*

Übungsaufgabe 6.3.13. *Eine Konservendose mit 850ml Inhalt soll so gebaut werden, dass der Blechverbrauch minimal ist. Wir vernachlässigen Materialstärke und Falznähte und nehmen an, dass es sich bei der Dose um einen einfachen Zylinder mit Radius r und Höhe h handelt. Welchen Radius und welche Höhe hat die optimale Dose? Vergleichen Sie Ihre gefundenen Werte mit den Maßen einer real existierenden Dose aus dem Supermarkt.*

Hinweis: *Es ist sinnvoll, zunächst Funktionen aufzustellen, die das Volumen und den Blechverbrauch für gegebenen Radius und Höhe beschreiben. Da das Volumen konstant 850ml sein soll, können Sie dann die Höhe als Funktion des Radius beschreiben.*

Übungsaufgabe 6.3.14. *Bestimmen Sie alle lokalen Extremstellen der Funktion*

$$f\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto x^2 \mathrm{e}^x - 2x\mathrm{e}^x - \frac{x^3}{3} + 2x + 5.$$

Übungsaufgabe 6.3.15. *Sei $f\colon \mathbb{R} \to \mathbb{R}$ eine stetig differenzierbare Funktion, die die Gleichung $f'(x) = \lambda f(x)$ für eine reelle Zahl λ erfüllt. Zeigen Sie $f(x) = a\exp(\lambda x)$ für eine reelle Zahl $a \in \mathbb{R}$.*

6.3.2. Konvexität und Wendepunkte.

Definition 6.3.16. *Sei $I \subseteq \mathbb{R}$ ein Intervall. Eine reell-wertige Funktion $f\colon I \to \mathbb{R}$ heißt* konvex, *wenn für alle $x_1, x_2 \in I$ und alle $\lambda \in (0, 1)$ gilt*

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

Eine Funktion $f\colon I \to \mathbb{R}$ heißt konkav, *wenn $-f$ konvex ist.*

Proposition 6.3.17. *Sei $f\colon I \to \mathbb{R}$ zweimal stetig differenzierbar auf einem offenen Intervall I. Dann gilt*

(1) f ist genau dann konvex auf I, wenn $f''(x) \geq 0$ für alle $x \in I$ gilt.
(2) f ist genau dann konkav auf I, wenn $f''(x) \leq 0$ für alle $x \in I$ gilt.

Beweis. Wir zeigen (1), (2) folgt durch Betrachtung von $-f$.

Sei $f''(x) \geq 0$ für alle $x \in I$. Dann ist $f'\colon D \to \mathbb{R}$ nach Korollar 6.3.5 monoton wachsend. Seien $x_1, x_2 \in I$ mit $x_1 < x_2$. Für $0 < \lambda < 1$ setzen wir $x = \lambda x_1 + (1 - \lambda)x_2$, und es gilt $x_1 < x < x_2$. Nach dem Mittelwertsatz 6.3.4 existieren $z_1 \in (x_1, x)$ und $z_2 \in (x, x_2)$ mit

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(z_1) \leq f'(z_2) = \frac{f(x_2) - f(x)}{x_2 - x}.$$

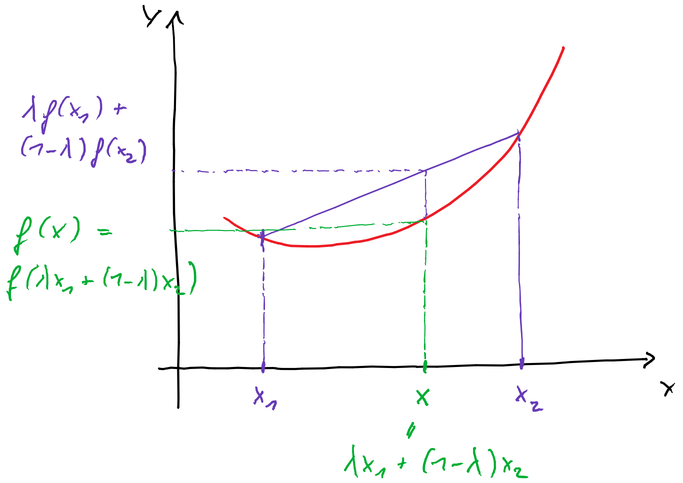


ABBILDUNG 21. Konvexität von f: Die Konvexitätsbedingung bedeutet, dass der Graph unterhalb der Sekante liegt, die $(x_1, f(x_1))$ und $(x_2, f(x_2))$ verbindet.

Mit $x - x_1 = (1-\lambda)(x_2 - x_1)$ und $x_2 - x = \lambda(x_2 - x_1)$ folgt

$$\frac{f(x) - f(x_1)}{1-\lambda} \leq \frac{f(x_2) - f(x)}{\lambda}.$$

Durch Umstellen folgt $f(x) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$, die Funktion ist also konvex.

Sei nun $f\colon I \to \mathbb{R}$ konvex. Wir nehmen an, dass $f''(x) \geq 0$ nicht für alle $x \in I$ gilt, es existiert also ein $x_0 \in I$ mit $f''(x_0) < 0$. Sei $c = f'(x_0)$. Wir definieren auf I die zweimal stetig differenzierbare Funktion $\phi(x) := f(x) - c(x - x_0)$. Nach Konstruktion gilt $\phi'(x_0) = 0$ und $\phi''(x_0) = f''(x_0) < 0$. Nach Korollar 6.3.7 hat ϕ in x_0 ein lokales Minimum, also existiert ein $\epsilon > 0$, so dass $[x_0 - \epsilon, x_0 + \epsilon] \subseteq I$ und $\phi(x_0 - \epsilon) < \phi(x_0)$ und $\phi(x_0 + \epsilon) < \phi(x_0)$ gilt. Daraus folgt

$$f(x_0) = \phi(x_0) > \frac{\phi(x_0 - \epsilon) + \phi(x_0 + \epsilon)}{2} = \frac{f(x_0 - \epsilon) + f(x_0 + \epsilon)}{2}.$$

Mit $x_1 = x_0 - \epsilon$, $x_2 = x_0 + \epsilon$ und $\lambda = \frac{1}{2}$ haben wir $x_0 = \lambda x_1 + (1-\lambda)x_2$. Dann haben wir

$$f(\lambda x_1 + (1-\lambda)x_2) > \lambda f(x_1) + (1-\lambda) f(x_2),$$

was im Widerspruch zur Konvexität von f steht. □

DEFINITION 6.3.18. *Für eine Funktion $f\colon I \to \mathbb{R}$ heißt $x_0 \in D$* Wendepunkt, *wenn ein $\epsilon > 0$ existiert, so dass f auf $(x_0 - \epsilon, x_0)$ konvex (bzw. konkav) ist und auf $(x_0, x_0 + \epsilon)$ konkav (bzw. konvex) ist.*

Wenn f zweimal stetig differenzierbar ist, dann x_0 ein Wendepunkt, wenn $f''(x_0) = 0$ und f'' am Punkt x_0 das Vorzeichen wechselt.

6.3.3. Kurvendiskussion.

(1) Bestimmung maximaler Definitionsbereich (existieren stetige Fortsetzungen?), Bestimmung Wertebereich
(2) Symmetrie: ist f eine gerade Funktion, d.h. $f(x) = f(-x)$, oder eine ungerade Funktion, d.h. $f(x) = -f(-x)$?
(3) ist f periodisch?
(4) Stetigkeit/Differenzierbarkeit von f
(5) Nullstellen

(6) Extremstellen von f, Intervalle, in denen f monoton ist.
(7) Wendepunkte von f, Intervalle, in denen f konvex bzw. konkav ist.
(8) Grenzwerte am Rand des Definitionsbereichs, asymptotische Entwicklung
(9) Skizze, basierend auf den relevanten Orientierungspunkten

ÜBUNGSAUFGABE 6.3.19. *Gegeben sei eine reelle Zahl a und die Funktion $f\colon \mathbb{R}_+ \to \mathbb{R}\colon x \mapsto x^a \exp(-x)$. Untersuchen Sie f hinsichtlich Monotonie, Konvexität und lokaler bzw. globaler Extrema.*

6.4. Anwendungen

6.4.1. Ableitung von Potenzreihen.

PROPOSITION 6.4.1. *Seien $f_n\colon I \to \mathbb{R}$, $n \in \mathbb{N}$, auf I differenzierbare Funktionen, so dass die folgenden Aussagen gelten:*

(1) Die Reihe $\sum_{n=0}^\infty f_n$ konvergiert punktweise auf I.
(2) Für jede Funktion $f_n'\colon I \to \mathbb{R}$ gilt für alle $x \in I$ eine Abschätzung $|f_n'(x)| \leq M_n$, so dass $\sum_{n=0}^\infty M_n$ konvergiert.

Dann ist $f = \sum_{n=0}^\infty f_n$ auf I differenzierbar und hat die Ableitung

$$f' = \sum_{n=0}^\infty f_n'.$$

BEWEIS. Für $\epsilon > 0$ wollen wir zeigen, dass ein $\delta > 0$ existiert, so dass für alle x mit $|x - x_0| < \delta$ gilt

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - \sum_{n=0}^\infty f_n'(x_0) \right| < \epsilon.$$

Für $x \in I \setminus \{x_0\}$ haben wir mit der Dreiecksungleichung und der Voraussetzung $\lim_{n\to\infty} f_n(x) = f(x)$:

$$\begin{aligned} \left| \frac{f(x) - f(x_0)}{x - x_0} - \sum_{n=0}^\infty f_n'(x_0) \right| &\leq \sum_{n=0}^N \left| \frac{f_n(x) - f_n(x_0)}{x - x_0} - f_n'(x_0) \right| \\ &+ \left| \sum_{n=N+1}^\infty \frac{f_n(x) - f_n(x_0)}{x - x_0} \right| \\ &+ \left| \sum_{n=N+1}^\infty f_n'(x_0) \right|. \end{aligned}$$

Aus der Voraussetzung (2) haben wir $|f_n'(x)| \leq M_n$. Aus dem Mittelwertsatz 6.3.4 folgt, dass für alle $x \in I$ gilt

$$\left| \frac{f_n(x) - f_n(x_0)}{x - x_0} \right| \leq M_n.$$

Damit haben wir

$$\left| \sum_{n=N+1}^\infty \frac{f_n(x) - f_n(x_0)}{x - x_0} \right| \leq \sum_{n=N+1}^\infty M_n.$$

Wegen der Konvergenz von $\sum_{n=0}^\infty M_n$ existiert ein N, so dass

$$\sum_{n=N+1}^\infty M_n < \frac{\epsilon}{3}$$

gilt, damit haben wir den zweiten Term abgeschätzt.

Außerdem folgt aus der Voraussetzung (2) (bzw. der daraus resultierenden gleichmäßigen Konvergenz von $\sum_{n=0}^{\infty} f_n'$), dass wir (nach eventuellem Vergrößern von N) den dritten Term auch entsprechend abschätzen können:

$$\left|\sum_{n=N+1}^{\infty} f_n'(x_0)\right| < \frac{\epsilon}{3}.$$

Für den ersten Term folgt aus der Differenzierbarkeit der f_n, dass es ein $\delta > 0$ gibt, so dass für alle $x \in I$ mit $|x - x_0| < \delta$ gilt

$$\sum_{n=0}^{N} \left|\frac{f_n(x) - f_n(x_0)}{x - x_0} - f_n'(x_0)\right| < \frac{\epsilon}{3}$$

Damit haben wir die gewünschte Abschätzung bewiesen. □

KOROLLAR 6.4.2. *Sei $f(x) = \sum_{k=0}^{\infty} a_k x^k$ eine Potenzreihe mit Konvergenzradius $R > 0$. Für alle $\rho \in (0, R)$ konvergiert die Reihe*

$$\sum_{k=1}^{\infty} k a_k x^{k-1}$$

absolut und gleichmäßig auf $|x| \leq \rho$ gegen f'.

BEWEIS. Die Ableitung von $f_n(x) = \sum_{k=0}^{n} a_k x^k$ ist $f_n'(x) = \sum_{k=1}^{n} k a_k x^{k-1}$. Für $|x| \leq \rho$ existiert ein q mit $\frac{\rho}{R} < q < 1$, so dass gilt

$$\left|k a_k x^{k-1}\right| < ckq^{k-1}.$$

Mit der geometrischen Reihe sind dann die Voraussetzungen für Proposition 6.4.1 erfüllt. □

BEISPIEL 6.4.3. *Aus Übungsaufgabe 6.2.8, kombiniert mit der geometrischen Reihe, folgt*

$$\arctan'(x) = \frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-1)^n x^{2n}.$$

Mit Korollar 6.4.2 sehen wir aber auch, dass die Potenzreihe die Ableitung der Potenzreihe

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1}$$

ist. Aus dem Mittelwertsatz folgt dann

$$\arctan x = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1} + c$$

für eine Konstante c, und durch Betrachtung von $x = 0$ sehen wir $c = 0$.

Aus $\arctan 1 = \frac{\pi}{4}$ *folgt dann*

$$\frac{\pi}{4} = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}$$

□

ÜBUNGSAUFGABE 6.4.4. *Gegeben sei die Potenzreihe*

$$\sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k}.$$

Bestimmen Sie den Konvergenzradius. Zeigen Sie, dass die dadurch dargestellte Funktion differenzierbar ist, und die Ableitung auf dem Konvergenzkreis durch die

Funktion $\frac{1}{1+x}$ *gegeben ist. Folgern Sie, dass die Potenzreihe die Funktion* $\ln(1+x)$ *darstellt. Folgern Sie daraus, dass*

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} = \ln(2).$$

6.4.2. Regeln von de l'Hospital.

SATZ 6.4.5. *Seien* $f, g\colon (a,b) \to \mathbb{R}$ *differenzierbar, und sei* $g'(x) \neq 0$ *für alle* $x \in (a,b)$. *Wenn*

$$\lim_{x\to b^-} f(x) = \lim_{x\to b^-} g(x) = 0, \qquad \textit{oder} \qquad \lim_{x\to b^-} f(x) = \lim_{x\to b^-} g(x) = \infty$$

und

$$\lim_{x\to b^-} \frac{f'(x)}{g'(x)} = c,$$

dann gilt

$$\lim_{x\to b^-} \frac{f(x)}{g(x)} = \lim_{x\to b^-} \frac{f'(x)}{g'(x)}$$

Analoge Aussagen gelten für die Fälle $x \to a^+$ *oder* $x \to \pm\infty$.

BEWEIS. Wir beweisen den Fall $\lim_{x\to b^-} f(x) = \lim_{x\to b^-} g(x) = 0$. In diesem Fall können wir f und g stetig durch 0 auf $(a,b]$ fortsetzen. Nach dem verallgemeinerten Mittelwertsatz 6.3.9 existiert für jedes $z \in (a,b)$ ein x_0 mit

$$\frac{f'(x_0)}{g'(x_0)} = \frac{f(b)-f(z)}{g(b)-g(z)}.$$

Für $z \to b^-$ gilt $x_0 \to b^-$ und wir haben

$$\lim_{z\to b^-} \frac{f(z)}{g(z)} = \lim_{x_0\to b^-} \frac{f'(x_0)}{g'(x_0)},$$

und der Grenzwert existiert nach Voraussetzung. □

BEISPIEL 6.4.6. *Bereits bekannte Grenzwerte, die auch mit den* L'HOSPITAL*schen Regeln ausgewertet werden können sind:*

$$\lim_{x\to 0} \frac{\sin x}{x} = 1, \qquad \lim_{x\to\infty} \frac{\log x}{x^n} = 0.$$

□

ÜBUNGSAUFGABE 6.4.7. *Berechnen Sie die folgenden Grenzwerte:*

$$\lim_{x\to 0} x\cot x, \qquad \lim_{x\to 0}\left(\frac{1}{\sin x} - \frac{1}{x}\right), \qquad \lim_{x\to\infty}\left(\sqrt{x+\sqrt{x}} - \sqrt{x}\right).$$

$$\lim_{x\to 0} \frac{\cos(x)-1}{x^2}, \qquad \lim_{x\to 1} \frac{\ln(x)-x+1}{x^2-2x+1}$$

ÜBUNGSAUFGABE 6.4.8. *Das* PLANCK*sche Strahlungsgesetz beschreibt die Energieverteilung der Wärmestrahlung eines schwarzen Körper in Abhängigkeit von der Wellenlänge* λ:

$$\frac{k_1}{\lambda^5} \frac{1}{\mathrm{e}^{\frac{k_2}{\lambda}} - 1}.$$

Hierbei sind die verschiedenen für die Betrachtung nicht relevanten Größen zu Konstanten k_1 *und* k_2 *zusammengefasst. Bestimmen Sie den Grenzwert für* $\lambda \to 0$.

6.4.3. Banachscher Fixpunktsatz.

DEFINITION 6.4.9. *Sei $f\colon [a,b] \in \mathbb{R}$ eine Abbildung. Ein Punkt $x \in [a,b]$ heißt* Fixpunkt *von f, wenn $f(x) = x$ gilt.*

SATZ 6.4.10 (BANACHscher Fixpunktsatz). *Sei $f\colon [a,b] \to \mathbb{R}$ eine stetig differenzierbare Funktion, so dass die beiden folgenden Aussagen gelten:*

(1) $a \leq f(x) \leq b$ für alle $x \in [a,b]$
(2) es existiert $K < 1$, mit $|f'(x)| \leq K$ für alle $x \in (a,b)$.

Dann gelten die folgenden Aussagen

(B1) Es existiert genau ein $p \in [a,b]$ mit $f(p) = p$.
(B2) Für alle $x_0 \in [a,b]$ konvergiert die Folge $x_{n+1} = f(x_n)$ gegen p.
(B3) Es gilt $|x_n - p| \leq \frac{K}{1-K}|x_n - x_{n-1}|$ für alle natürlichen $n \geq 1$.

BEWEIS. Wir wenden den Zwischenwertsatz 4.4.1 auf die Funktion $f(x) - x$ an. Nach Voraussetzung (1) ist $f(a) - a \geq 0$ und $f(b) - b \leq 0$, also existiert eine Nullstelle, damit haben wir die Existenz in (B1). Für die Eindeutigkeit in (B1) seien $x_1 < x_2$ zwei Fixpunkte. Nach dem Mittelwertsatz 6.3.4 existiert ein $x_0 \in (x_1, x_2)$ mit

$$f'(x_0) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} = \frac{x_2 - x_1}{x_2 - x_1} = 1.$$

Dies ist ein Widerspruch zur Voraussetzung $f'(x) \leq K < 1$ für alle $x \in (a,b)$.

Sei p der Fixpunkt, $x_0 \in [a,b]$ beliebig. Wir betrachten die durch $x_{n+1} = f(x_n)$ definierte Folge. Nach Definition der Folge bzw. der Fixpunkteigenschaft von p gilt $x_n - p = f(x_{n-1}) - f(p)$. Aus dem Mittelwertsatz 6.3.4 existiert ein y zwischen x_{n-1} und p, so dass

$$f'(y) = \frac{f(x_{n-1}) - f(p)}{x_{n-1} - p}.$$

Dann ist

$$|x_n - p| = |f(x_{n-1}) - f(p)| = |f'(y)||x_{n-1} - p| \leq K|x_{n-1} - p|$$

Induktiv folgt $|x_n - p| \leq K^n|x_0 - p|$, und aus $\lim_{n\to\infty} K^n = 0$ sehen wir $\lim_{n\to\infty} x_n = p$.

Für (B3) haben wir mit der Dreiecksungleichung

$$|x_n - p| \leq K|x_{n-1} - p| \leq K|x_{n-1} - x_n| + K|x_n - p|.$$

Umstellen liefert die Behauptung. □

6.4.4. Newton-Verfahren zur Bestimmung von Nullstellen. Zur numerischen/näherungsweisen Bestimmung von Nullstellen von Funktionen können wir das Newton-Verfahren anwenden. Das Verfahren ergibt sich als Spezialfall aus dem BANACHschen Fixpunktsatz.

Sei dafür $f\colon I \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion auf dem Intervall I. Für einen Startwert x_0 hat die Tangente an den Funktionsgraphen im Punkt x_0 mit die Gleichung $y = f(x_0) + f'(x_0)(x - x_0)$. Der Schnittpunkt von Tangente und x-Achse erfüllt $f(x_0) + f'(x_0)(x_1 - x_0) = 0$, ist also durch

$$x_1 = \frac{x_0 f'(x_0) - f(x_0)}{f'(x_0)} = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

gegeben. Dafür ist natürlich $f'(x_0) \neq 0$ Voraussetzung. Der Punkt x_1 ist eine (hoffentlich) bessere Näherung an die gesuchte Nullstelle von f.

Wir wenden in dieser Situation den BANACHschen Fixpunktsatz an. Dafür betrachten wir die Funktion

$$F(x) = x - \frac{f(x)}{f'(x)}.$$

Nach Definition ist dann $f(x) = 0$ genau dann, wenn x ein Fixpunkt von $F(x)$ ist. Es gilt

$$F'(x) = 1 - \frac{(f'(x))^2 - f(x)f''(x)}{(f'(x))^2} = \frac{f(x)f''(x)}{(f'(x))^2}$$

PROPOSITION 6.4.11. *Sei $f\colon [a,b] \to \mathbb{R}$ zweimal stetig differenzierbar, so dass die folgenden Aussagen gelten:*

(i) *f hat eine Nullstelle $z \in [a,b]$.*
(ii) *$f'(x) \neq 0$ für alle $x \in [a,b]$.*
(iii) *f ist in $[a,b]$ konkav oder konvex.*
(iv) *Die Iterationswerte $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$ liegen für $x_0 = a$ und $x_0 = b$ wieder in $[a,b]$.*

Dann konvergiert die rekursiv definierte Folge $(x_n)_n$ mit beliebigem Anfangswert $x_0 \in [a,b]$ und

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

gegen die Nullstelle $z \in [a,b]$. Für $m = \min\{|f'(x)| \mid x \in [a,b]\}$ und $M = \max\{|f''(x)| \mid x \in [a,b]\}$ gilt

$$|z - x_n| \leq \frac{M}{2m}|x_n - x_{n-1}|^2.$$

ÜBUNGSAUFGABE 6.4.12. *Zeigen Sie, dass das* NEWTON-*Verfahren für $f(x) = x^2 - 2$ für alle $x_0 \neq 0$ gegen eine Nullstelle konvergiert.*

ÜBUNGSAUFGABE 6.4.13. *Benutzen Sie das Newton-Verfahren, um* $\ln(2)$ *auf 4 Nachkommastellen genau zu berechnen.*

Hinweis: *Bestimmen Sie die Nullstelle von* $\exp(x) - 2$ *mit dem Verfahren aus der Vorlesung. Überprüfen Sie die Voraussetzungen des* BANACH*schen Fixpunktsatzes für die Hilfsfunktion! Als linker Intervallrand und Startwert kann $\frac{1}{2}$ benutzt werden (warum?). Wie genau ist der Wert tatsächlich, wenn das Verfahren abbricht? Für die Berechnung von Werten der Exponentialfunktion kann Taschenrechner bzw. Computeralgebrasystem benutzt werden. Diskutieren Sie, welche Genauigkeit für diese Berechnungen notwendig ist (und wo das eingeht).*

ÜBUNGSAUFGABE 6.4.14. *Lösen Sie die Gleichung $x = \cos x$ mit Hilfe des Newton-Verfahrens.*

ÜBUNGSAUFGABE 6.4.15. *Finden Sie eine Stelle $x_0 \in \mathbb{R}$, an der das* NEWTON-*Verfahren für $f(x) = x^3 - 5x$ nicht konvergiert. (mit Begründung)*

KAPITEL 7

Integralrechnung in einer Variablen

Integrale sind ein wesentliches Hilfsmittel zur Berechnung von Längen, Flächen, Volumina (oder in der Physik geleistete Arbeit, zurückgelegter Weg,...). Im einfachsten Fall berechnet das bestimmte Integral die Fläche unter dem Graphen einer Funktion, indem die Fläche durch einfachere Flächen (z.B. Vereinigung von Rechtecken) angenähert wird.

7.1. Regelfunktionen und Approximation durch Treppenfunktionen

DEFINITION 7.1.1. *Sei $[a,b]$ ein Intervall. Eine Menge von Punkten $x_0, \dots, x_n$ mit*

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

heißt Zerlegung des Intervalls $[a,b]$. *Eine Funktion $\phi\colon [a,b] \to \mathbb{R}$ heißt* Treppenfunktion *bzw.* stückweise konstante Funktion, *wenn es eine Zerlegung $a = x_0, \dots, x_n = b$ des Intervalls $[a,b]$ gibt, so dass ϕ in jedem offenen Intervall (x_i, x_{i+1}) konstant ist. Dabei können die Funktionswerte $f(x_i)$ beliebig sein.*

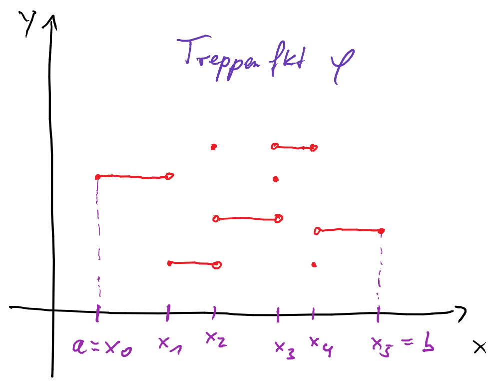


ABBILDUNG 22. Beispiel einer Treppenfunktion

DEFINITION 7.1.2 (Regelfunktion). *Eine Funktion $f\colon [a,b] \to \mathbb{R}$ heißt* Regelfunktion auf $[a,b]$, *wenn*

(1) in jedem Punkt $z \in [a,b)$ ein rechtsseitiger Grenzwert $\lim_{x\to z^+} f(x)$ existiert, und
(2) in jedem Punkt $z \in (a,b]$ ein linksseitiger Grenzwert $\lim_{x\to z^-} f(x)$ existiert.

DEFINITION 7.1.3. *Eine beschränkte Funktion $f\colon [a,b] \to \mathbb{R}$ heißt* stückweise stetig, *wenn es eine Zerlegung $Z = (x_0, x_1, \dots, x_n)$ und stetige Funktionen*

$$g_i\colon [x_i, x_{i+1}] \to \mathbb{R}$$

gibt, so dass für alle $0 \le i < n$ gilt $g_i|_{(x_i,x_{i+1})} = f|_{(x_i,x_{i+1})}$.

Beispiel 7.1.4. *Stückweise stetige Funktionen sind Regelfunktionen.* □

Übungsaufgabe 7.1.5. *Zeigen Sie, dass monotone Funktionen Regelfunktionen sind.*

Hinweis: *Überlegen Sie sich, warum monotone Funktionen auf abgeschlossenen Intervallen beschränkt sein müssen.*

Satz 7.1.6 (Approximationssatz). *Eine Funktion* $f\colon [a,b] \to \mathbb{R}$ *ist genau dann eine Regelfunktion, wenn es eine gleichmäßig konvergente Folge* $(\phi_n)_n$ *von Treppenfunktionen mit* $\lim_{n\to\infty} \phi_n = f$ *gibt.*

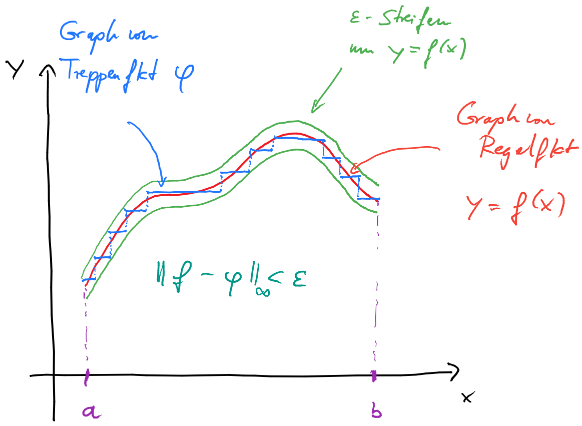


Abbildung 23. Approximation von Regelfunktionen durch Treppenfunktionen.

Beweis. Für die Aussage müssen wir zeigen, dass es zu jedem $\epsilon > 0$ eine Treppenfunktion $\phi\colon [a,b] \to \mathbb{R}$ gibt, so dass $\|f-\phi\|_\infty \le \epsilon$ gilt, d.h. $|f(x)-\phi(x)| \le \epsilon$ für alle $x \in [a,b]$.

"$\Rightarrow$": Widerspruchsbeweis. Sei f eine Regelfunktion. Wir nehmen an, dass die ein $\epsilon > 0$ existiert, so dass es keine Treppenfunktion ϕ mit $\|f-\phi\|_\infty \le \epsilon$ gibt. Wir konstruieren rekursiv eine Intervallschachtelung $I_k = [a_k, b_k]$, so dass auf keinem der Intervalle $[a_k, b_k]$ eine Treppenfunktion ϕ mit $\|f-\phi\|_\infty \le \epsilon$ existiert. Der Rekursionsanfang ist $[a_1, b_1] = [a,b]$. Haben wir das Intervall $[a_k, b_k]$ gegeben, halbieren wir wieder durch den Mittelpunkt $m = \frac{a_k+b_k}{2}$. Dann gilt für eines der beiden Intervalle $[a_k, m]$ und $[m, b_k]$, dass keine ϵ-approximierende Treppenfunktion für f existiert (sonst würde eine solche Treppenfunktion auf $[a_k, b_k]$ existieren); das entsprechende halbe Intervall ist jetzt das nächste Intervall $[a_{k+1}, b_{k+1}]$ in der Intervallschachtelung.

Sei z der in allen Intervallen $[a_k, b_k]$ liegende Punkt. Wir betrachten den Fall $z \in (a,b)$. Die Fälle $z = a$ und $z = b$ folgen analog. Nach Definition existieren die links- und rechtsseitigen Grenzwerte $\lim_{x\to z^+} f(x) = \lambda_+$ und $\lim_{x\to z^-} f(x) = \lambda_-$. Insbesondere existiert auch eine Zahl δ, so dass

$$|f(x)-\lambda_-| < \epsilon \text{ für } x \in [z-\delta, z), \qquad |f(x)-\lambda_+| < \epsilon \text{ für } x \in (z, z+\delta]$$

gilt. Auf $[z-\delta, z+\delta]$ haben wir eine Treppenfunktion

$$\phi(x) = \begin{cases} \lambda_- & x \in [z-\delta, z) \\ f(z) & x = z \\ \lambda_+ & (z, z+\delta], \end{cases}$$

die f auf $[z-\delta, z+\delta]$ (und damit auf jedem Intervall $[a_k, b_k] \subseteq [z-\delta, z+\delta]$) ϵ-gut approximiert. Das ist ein Widerspruch.

"$\Leftarrow$": Wir nehmen an, dass f gleichmäßiger Grenzwert von Treppenfunktionen ist und beweisen die Existenz des rechtsseitigen Grenzwerts für beliebiges $z \in [a, b)$. Für $\epsilon > 0$ sei ϕ eine Treppenfunktion mit $\|f - \phi\|_\infty \leq \frac{\epsilon}{2}$ und sei (z, β) ein Intervall, auf dem ϕ konstant ist. Dann gilt für beliebige Punkte $x, y \in (z, \beta)$

$$|f(x) - f(y)| \leq |f(x) - \phi(x)| + |\phi(y) - f(y)| \leq \epsilon.$$

Daraus folgt die Existenz des rechtsseitigen Grenzwerts: für eine Folge $(x_n)_n$ mit $\lim_{n\to\infty} x_n = z$ gilt für alle $x_n, x_m \in (z, \beta)$ schon $|f(x_n) - f(x_m)| \leq \epsilon$, die Folge konvergiert also nach dem Cauchy-Kriterium. Genauso sehen wir auch, dass der Grenzwert unabhängig von der Folge $(x_n)_n$ mit $x_n \to z$ ist. □

Bemerkung 7.1.7. *Vergleich: reelle Zahlen als Grenzwerte von Folgen rationaler Zahlen. Hier erhalten wir interessante Funktionen, indem wir gleichmäßige Grenzwerte von sehr einfachen Funktionen betrachten.*

Übungsaufgabe 7.1.8. *Zeigen Sie, dass Regelfunktionen beschränkt sind.*

Bemerkung 7.1.9. *Man kann zeigen, dass Regelfunktionen nur höchstens abzählbar viele Unstetigkeitsstellen haben.*

7.2. Regelintegral und Riemann-Summen

Wir definieren erst das Integral für Treppenfunktionen und setzen diese Definition dann durch einen Grenzübergang auf Regelfunktionen fort.

Definition 7.2.1. *Sei $\phi\colon [a, b] \to \mathbb{R}$ eine Treppenfunktion, so dass eine Zerlegung $a = x_0, \ldots, x_n = b$ von $[a, b]$ existiert und ϕ auf den Intervallen (x_i, x_{i+1}) jeweils den konstanten Wert c_i hat. Dann definieren wir das Integral*

$$\int_a^b \phi(x)\mathrm{d}x := \sum_{i=0}^{n-1} c_i(x_{i+1} - x_i).$$

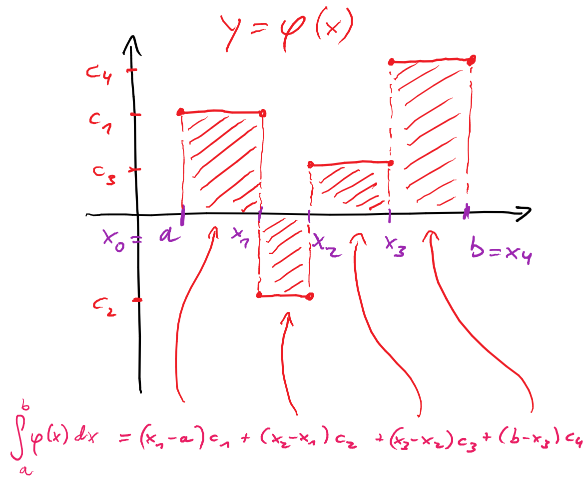


Abbildung 24. Das Integral einer Treppenfunktion ist die Summe der Rechteckflächen zwischen Graph und x-Achse.

Lemma 7.2.2. *Diese Definition ist unabhängig von der gewählten Zerlegung.*

Beweis. Für zwei Zerlegungen $Z = \{a = x_0, x_1, \ldots, x_n = b\}$ und $Z' = \{a = y_0, y_1, \ldots, y_m = b\}$ ist $Z \cup Z'$ auch wieder eine Zerlegung von $[a, b]$. Induktiv reicht es zu zeigen, dass sich das Integral nicht verändert, wenn eine Zerlegung $a = x_0, \ldots, x_n = b$ durch Einfügen eines neuen Punkts t zwischen x_k und x_{k+1} verfeinert wird. In diesem Fall ändert sich nur der Summand $c_k(x_{k+1} - x_k)$ zu $c_k(x_{k+1} - t) + c_k(t - x_k)$, also bleibt das Integral gleich. □

Proposition 7.2.3. *Für Treppenfunktionen $\phi, \psi\colon [a, b] \to \mathbb{R}$ gelten die folgenden Eigenschaften:*

Linearität: *Für $\lambda, \mu \in \mathbb{R}$ gilt*

$$\int_a^b (\lambda\phi(x) + \mu\psi(x))\mathrm{d}x = \lambda \int_a^b \phi(x)\mathrm{d}x + \mu \int_a^b \psi(x)\mathrm{d}x.$$

Beschränktheit:

$$\left| \int_a^b \phi(x)\mathrm{d}x \right| \leq \int_a^b |\phi(x)|\mathrm{d}x \leq (b - a)\|\phi\|_\infty.$$

Monotonie: *Wenn $\phi(x) \leq \psi(x)$ für alle $x \in [a, b]$, dann gilt*

$$\int_a^b \phi(x)\mathrm{d}x \leq \int_a^b \psi(x)\mathrm{d}x.$$

Proposition 7.2.4. *Sei $f\colon [a, b] \to \mathbb{R}$ eine Regelfunktion. Für jede gleichmäßig konvergente Folge $(\phi_n)_n$ von Treppenfunktionen auf $[a, b]$ mit $\lim_{n\to\infty} \phi_n = f$ existiert der Grenzwert*

$$\lim_{n\to\infty} \int_a^b \phi_n(x)\mathrm{d}x.$$

Der Grenzwert hängt nicht von der Wahl der Folge von Treppenfunktionen ab.

Beweis. Für die Existenz des Grenzwerts zeigen wir, dass die Zahlen $I_n = \int_a^b \phi_n(x)\mathrm{d}x$ eine Cauchy-Folge bilden. Aus der Beschränktheit des Integrals für Treppenfunktionen folgt

$$|I_n - I_m| \leq (b - a)\|\phi_n - \phi_m\|_\infty \leq (b - a)(\|\phi_n - f\|_\infty - \|f - \phi_m\|_\infty).$$

Nach Voraussetzung ist $(\|f - \phi_k\|)_k$ eine Nullfolge, damit ist $(I_n)_n$ eine Cauchy-Folge, also konvergent.

Für die Unabhängigkeit sei $(\psi_n)_n$ eine weitere Folge von Treppenfunktionen auf $[a, b]$, die gleichmäßig gegen f konvergiert. Wir erhalten durch $\phi_0, \psi_0, \phi_1, \psi_1, \ldots$ eine Folge (χ_n) von Treppenfunktionen, die gleichmäßig gegen f konvergiert. Nach dem Existenzteil der Aussage konvergiert die Folge $\left(\int_a^b \chi_n(x)\mathrm{d}x\right)_n$, und der Grenzwert stimmt dann mit den Grenzwerten der beiden Teilfolgen $\left(\int_a^b \phi_n(x)\mathrm{d}x\right)_n$ und $\left(\int_a^b \psi_n(x)\mathrm{d}x\right)_n$ überein. □

Definition 7.2.5. *Sei $f\colon [a, b] \to \mathbb{R}$ eine Regelfunktion. Wir definieren das* bestimmte Integral *von f auf $[a, b]$ durch*

$$\int_a^b f(x)\mathrm{d}x := \lim_{n\to\infty} \int_a^b \phi_n(x)\mathrm{d}x,$$

wobei $(\phi_n)_n$ eine gleichmäßig konvergente Folge von Treppenfunktionen auf $[a, b]$ mit $\lim_{n\to\infty} \phi_n = f$ ist.

Bemerkung 7.2.6. *Konvention*

$$\int_a^a f(x)\mathrm{d}x = 0, \qquad \int_b^a f(x)\mathrm{d}x = -\int_a^b f(x)\mathrm{d}x.$$

Das Regelintegral kann durch Riemann-Summen approximiert werden.

Definition 7.2.7. *Gegeben eine Funktion* $f\colon [a,b] \to \mathbb{R}$ *und eine Zerlegung* $Z = \{a = x_0, x_1, \dots, x_n = b\}$ *von* $[a,b]$. *Für vorgegebene Stellen* $\xi_i \in [x_i, x_{i+1}]$ *heißt*

$$\sum_{i=0}^{n-1} f(\xi_i)(x_{i+1} - x_i)$$

Riemann-Summe *für Zerlegung die* Z. *Die* ξ_i *heißen* Stützstellen, *die* $f(\xi_i)$ Stützwerte, *und* $\max\{x_{i+1} - x_i \mid 0 \le i \le n-1\}$ *die* Feinheit *der Zerlegung* Z.

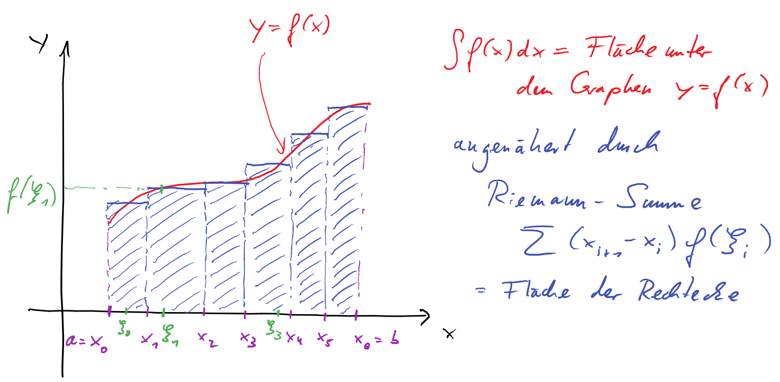


Abbildung 25. Annäherung des Integrals durch Riemann-Summe für Zerlegung des Intervalls $[a,b]$.

Definition 7.2.8. *Eine Funktion* $f\colon [a,b] \to \mathbb{R}$ *heißt* Riemann-integrierbar, *wenn für jede Folge* Z_n *von Zerlegungen, deren Feinheit gegen* 0 *konvergiert, und jede Wahl von Stützstellen* $\xi_i \in [x_i, x_{i+1}]$ *gilt*

$$\lim_{n\to\infty} \sum_{i=0}^{n-1} f(\xi_i)(x_{i+1} - x_i)$$

Die Zahl I *heißt dann* bestimmtes (Riemann-)Integral *von* f *auf* $[a,b]$.

Notation: $I = \int_a^b f(x)\mathrm{d}x$

Satz 7.2.9. *Sei* $f\colon [a,b] \to \mathbb{R}$ *eine Regelfunktion. Dann existiert zu jedem* $\epsilon > 0$ *ein* $\delta > 0$, *so dass für jede Zerlegung* $Z = \{a = x_0, \dots, x_n = b\}$ *von* $[a,b]$ *mit Feinheit* $\le \delta$ *und jede Wahl von Stützstellen* $\xi_i \in [x_i, x_{i+1}]$ *gilt*

$$\left|\sum_{i=0}^{n-1} f(\xi_i)(x_{i+1} - x_i) - \int_a^b f(x)\mathrm{d}x\right| \le \epsilon.$$

Insbesondere sind Regelfunktionen Riemann-integrierbar.

Beweis. Für Treppenfunktionen ist die Aussage leicht durch Induktion über die Anzahl der Sprungstellen zu beweisen. Für eine Regelfunktion f wählen wir eine Treppenfunktion ϕ mit $\|f - \phi\|_\infty < \frac{\epsilon}{3(b-a)}$. Dann existiert ein δ, so dass

$$\left|\sum_{k=0}^{n-1} \phi(\xi_k)(x_{k+1} - x_k) - \int_a^b \phi\mathrm{d}x\right| \le \frac{\epsilon}{3}.$$

Dann haben wir

$$\begin{aligned}
&\left|\sum_{k=0}^{n-1} f(\xi_k)(x_{k+1}-x_k) - \int_a^b f\mathrm{d}x\right| \\
\leq\ &\left|\sum_{k=0}^{n-1} f(\xi_k)(x_{k+1}-x_k) - \sum_{k=0}^{n-1} \phi(\xi_k)(x_{k+1}-x_k)\right| \\
+\ &\left|\sum_{k=0}^{n-1} \phi(\xi_k)(x_{k+1}-x_k) - \int_a^b \phi\mathrm{d}x\right| + \left|\int_a^b \phi\mathrm{d}x - \int_a^b f\mathrm{d}x\right| \\
\leq\ &\sum_{k=0}^{n-1} \|f-\phi\|_\infty (x_{k+1}-x_k) + \frac{\epsilon}{3} + \int_a^b \|f-\phi\|_\infty \mathrm{d}x \leq \epsilon.
\end{aligned}$$

□

Beispiel 7.2.10. *Wir betrachten die Dirichlet-Funktion*

$$f\colon [0,1] \to \mathbb{R}\colon x \mapsto \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

Wählen wir für eine Riemann-Summe als Stützstellen rationale Zahlen ξ_i, hat die Riemann-Summe Wert 1. Wählen wir als Stützstellen irrationale Zahlen, hat die Riemann-Summe Wert 0. Dies ist unabhängig von der Feinheit der Zerlegung, da jedes Intervall (x_i, x_{i+1}) sowohl rationale als auch irrationale Zahlen enthält. Damit ist die Dirichlet-Funktion nicht Riemann-integrierbar. □

Übungsaufgabe 7.2.11. *Berechnen Sie die folgenden bestimmten Integrale mittels Riemann-Summen:*

$$\int_0^a x\mathrm{d}x \text{ (für } a > 0), \qquad \int_0^a \cos(x)\mathrm{d}x \text{ (für } a > 0), \qquad \int_1^a \frac{\mathrm{d}x}{x} \text{ (für } a > 1).$$

Übungsaufgabe 7.2.12. *Wir betrachten die Funktion $f\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto x^2$. Berechnen Sie das Integral $\int_0^a f(x)\mathrm{d}x$ durch den Grenzwert geeigneter Riemann-Summen.*

(a) Zeigen Sie, dass für alle natürlichen Zahlen $n \geq 1$ gilt

$$\sum_{i=1}^{n} i^2 = 1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

(b) Für eine gegebene natürliche Zahl $n \geq 1$ betrachten wir die Zerlegung $Z_n = (0 = x_0, x_1, \ldots, x_n = a)$ von $[0,a]$ mit $x_k = \frac{k}{n}a$. Berechnen Sie die Riemann-Summe

$$S_{Z_n} = \sum_{i=0}^{n-1} f(\xi)(x_{i+1} - x_i)$$

mit den Stützstellen $\xi = x_i$. Zeigen Sie, dass der Grenzwert $\lim_{n\to\infty} S_{Z_n}$ existiert und bestimmen Sie den Grenzwert.

(c) Begründen Sie, warum der Grenzwert aus (b) mit $\int_0^a f(x)\mathrm{d}x$ übereinstimmt.

7.3. Integrationsregeln und Mittelwertsatz

Proposition 7.3.1. *Seien $f, g\colon [a,b] \to \mathbb{R}$ Regelfunktionen. Dann gelten die folgenden Aussagen:*

Linearität: *Für $\lambda, \mu \in \mathbb{R}$ gilt*

$$\int_a^b (\lambda f(x) + \mu g(x))\,\mathrm{d}x = \lambda \int_a^b f(x)\mathrm{d}x + \mu \int_a^b g(x)\mathrm{d}x.$$

Additivität: *Für $c \in [a,b]$ gilt*

$$\int_a^b f(x)\mathrm{d}x = \int_a^c f(x)\mathrm{d}x + \int_c^b f(x)\mathrm{d}x.$$

Monotonie: *Wenn $f(x) \leq g(x)$ für alle $x \in [a,b]$, dann gilt*

$$\int_a^b f(x)\mathrm{d}x \leq \int_a^b g(x)\mathrm{d}x.$$

Wenn $m \leq f(x) \leq M$ für alle $x \in [a,b]$, dann gilt

$$m(b-a) \leq \int_a^b f(x)\mathrm{d}x \leq M(b-a).$$

Beschränktheit: *Die Funktion $|f|$ ist eine Regelfunktion und es gilt*

$$\left|\int_a^b f(x)\mathrm{d}x\right| \leq \int_a^b |f(x)|\mathrm{d}x \leq (b-a)\|f\|_\infty.$$

Beweis. Seien $(\phi_n)_n$ und $(\psi_n)_n$ Folgen von Treppenfunktionen, die gleichmäßig gegen f bzw. g konvergieren.

(1) Es gilt $\lim_{n\to\infty} \|(\lambda f + \mu g) - (\lambda\phi_n + \mu\psi_n)\|_\infty = 0$. Aus den Eigenschaften des Integrals für Treppenfunktionen, s. Proposition 7.2.3, folgt

$$\int_a^b (\lambda f + \mu g)\mathrm{d}x = \lim_{n\to\infty} \int_a^b (\lambda\phi_n + \mu\psi_n)\mathrm{d}x = \lambda \int_a^b f\mathrm{d}x + \mu \int_a^b g\mathrm{d}x.$$

(2) Die Aussage gilt offensichtlich für das Integral von Treppenfunktionen. Wenn $\phi\colon [a,c] \to \mathbb{R}$ eine Treppenfunktion mit $\|f-\phi\|_\infty < \epsilon$ auf $[a,c]$ und $\psi\colon [c,b] \to \mathbb{R}$ eine Treppenfunktion mit $\|f-\psi\|_\infty < \epsilon$ auf $[c,b]$ ist, dann ist

$$\sigma(x) := \begin{cases} \phi(x) & x \in [a,c] \\ \psi(x) & x \in (c,b] \end{cases}$$

eine Treppenfunktion mit $[a,b]$ mit $\|f-\sigma\|_\infty < \epsilon$. Damit folgt die Aussage für Regelfunktionen.

(3) Wir definieren $\phi_n^- := \phi_n - \|f-\phi_n\|_\infty$ und $\psi_n^+ := \psi_n + \|g-\psi_n\|_\infty$. Dann haben wir

$$\phi_n^- \leq f \leq g \leq \psi_n^+, \qquad \lim_{n\to\infty} \|f-\phi_n^-\|_\infty = 0, \qquad \lim_{n\to\infty} \|g-\psi_n^+\|_\infty = 0$$

Damit folgt aus der Monotonie des Integrals für Treppenfunktionen, s. Proposition 7.2.3:

$$\int_a^b f\mathrm{d}x = \lim_{n\to\infty} \int_a^b \phi_n^-\mathrm{d}x \leq \lim_{n\to\infty} \int_a^b \psi_n^+\mathrm{d}x = \int_a^b g\mathrm{d}x.$$

(4) Es gilt $\lim_{n\to\infty} \||f| - |\phi_n|\|_\infty = 0$ und damit

$$\int_a^b |f|\mathrm{d}x = \lim_{n\to\infty} \int_a^b |\phi_n|\mathrm{d}x.$$

Damit folgt die Behauptung aus der Beschränktheit des Integrals für Treppenfunktionen, s. Proposition 7.2.3:

$$\left|\int_a^b f\mathrm{d}x\right| = \lim_{n\to\infty} \left|\int_a^b \phi_n\mathrm{d}x\right| \leq \lim_{n\to\infty} \|\phi_n\|_\infty(b-a) = \|f\|_\infty(b-a).$$

□

Satz 7.3.2. *Sei $f\colon [a,b] \to \mathbb{R}$ stetig und $g\colon [a,b] \to \mathbb{R}$ eine Regelfunktion mit $g(x) \geq 0$ für alle $x \in [a,b]$. Dann existiert ein $\xi \in [a,b]$, so dass*

$$\int_a^b f(x)g(x)\mathrm{d}x = f(\xi) \int_a^b g(x)\mathrm{d}x$$

Beweis. Seien m bzw. M das Minimum bzw. Maximum von f auf $[a, b]$ (diese existieren nach Satz 4.4.7). Dann gilt für alle $x \in [a, b]$ die Abschätzung $mg(x) \leq f(x)g(x) \leq Mg(x)$. Daraus folgt

$$m \int_a^b g(x)\mathrm{d}x \leq \int_a^b f(x)g(x)\mathrm{d}x \leq M \int_a^b g(x)\mathrm{d}x.$$

Das bedeutet, dass ein $c \in [m, M]$ existiert, so dass

$$\int_a^b f(x)g(x)\mathrm{d}x = c \int_a^b g(x)\mathrm{d}x.$$

Nach dem Zwischenwertsatz 4.4.1 existiert ein $\xi \in [a, b]$ mit $c = f(\xi)$, daraus folgt die Behauptung. □

Bemerkung 7.3.3. *Im Spezialfall $g(x) = 1$ haben wir damit die Existenz von $\xi \in [a, b]$ mit*

$$\int_a^b f(x)\mathrm{d}x = f(\xi)(b - a)$$

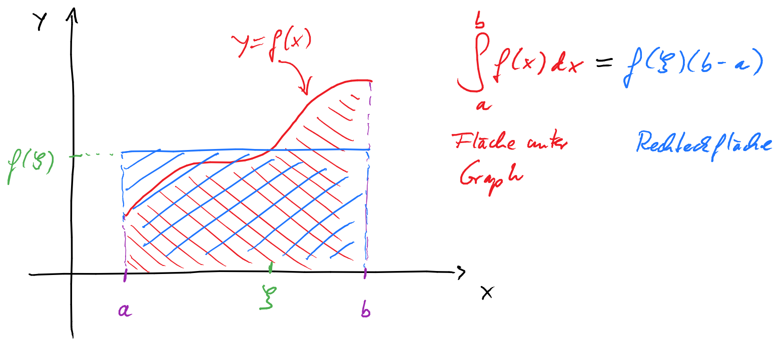


Abbildung 26. Mittelwertsatz der Integralrechnung: es existiert ein Punkt ξ, so dass das Integral gleich der Rechteckfläche $f(\xi)(b-a)$ ist.

7.4. Unbestimmtes Integral und Hauptsatz der Differential- und Integralrechnung

Definition 7.4.1. *Sei $f: I \to \mathbb{R}$ eine Funktion. Eine differenzierbare Funktion $F: I \to \mathbb{R}$ heißt* Stammfunktion *von f, wenn $F'(x) = f(x)$ für alle $x \in I$ ist.*

Die Menge aller Stammfunktionen von f wird mit $\int f(x)\mathrm{d}x$ bezeichnet und heißt unbestimmtes Integral *von f.*

Beispiel 7.4.2. *(1)* $\int x^a \mathrm{d}x = \frac{x^{a+1}}{a+1} + C$ *für* $a \neq -1$
(2) $\int \frac{1}{x}\mathrm{d}x = \ln|x| + C$
(3) $\int \mathrm{e}^{\lambda x}\mathrm{d}x = \frac{1}{\lambda}\mathrm{e}^{\lambda x} + C$ *für* $\lambda \neq 0$
(4) $\int \sin(x)\mathrm{d}x = -\cos(x) + C$
(5) $\int \cos(x)\mathrm{d}x = \sin(x) + C$
(6) $\int \frac{1}{1+x^2}\mathrm{d}x = \arctan(x) + C$
(7) $\int \frac{1}{\sqrt{1-x^2}}\mathrm{d}x = \arcsin(x) + C$ *für* $x \in (-1, 1)$
(8) $\int \frac{f'(x)}{f(x)}\mathrm{d}x = \ln|f(x)| + C$

□

Satz 7.4.3. *Sei $f: I \to \mathbb{R}$ stetig.*

(1) Die Funktion $F_a(x) = \int_a^x f(t)\mathrm{d}t$ ist eine Stammfunktion für f, d.h.

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_a^x f(t)\mathrm{d}t\right) = f(x).$$

Jede andere Stammfunktion von f hat die Form $F(x) = F_a(x) + c$ für geeignetes $c \in \mathbb{R}$.

(2) Für eine Stammfunktion F von f gilt

$$\int_a^b f(x)\mathrm{d}x = F(x)|_a^b = F(b) - F(a).$$

Beweis. (1) Nach den Rechenregeln für Integrale und dem Mittelwertsatz haben wir

$$\begin{aligned} F_a(x+h) - F_a(x) &= \int_a^{x+h} f(t)\mathrm{d}t - \int_a^x f(t)\mathrm{d}t \\ &= \int_x^{x+h} f(t)\mathrm{d}t \\ &= f(\xi_h)\cdot h \end{aligned}$$

für ein geeignetes $\xi_h \in [x, x+h]$. Dann ist $\lim_{h\to 0}\xi_h = x$, und aus der Stetigkeit von f folgt

$$F_a'(x) = \lim_{h\to 0}\frac{F_a(x+h) - F_a(x)}{h} = \lim_{h\to 0} f(\xi_h) = f(x).$$

Für eine andere Stammfunktion F ist $(F - F_a)' = F' - F_a' = f - f = 0$. Nach dem Mittelwertsatz der Differentialrechnung, Satz 6.3.4, ist $F - F_a$ konstant.

(2) Nach (1) ist $F_a = F - c$, und wir erhalten mit der Definition von F_a

$$\int_a^b f(x)\mathrm{d}x = F_a(b) - F_a(a) = F(b) - c - (F(a) - c) = F(b) - F(a).$$

□

Bemerkung 7.4.4. *Die Stammfunktion kann also zur Berechnung bestimmter Integrale benutzt werden. Wenn F eine Stammfunktion von f ist, wird folgende Notation häufig benutzt:*

$$\int_a^b f(t)\mathrm{d}t = F(x)|_a^b = F(b) - F(a)$$

7.5. Integrationstechniken

7.5.1. Partielle Integration. Aus der Produktregel folgt, dass für stetig differenzierbare Funktionen $u, v\colon I \to \mathbb{R}$ das Produkt uv eine Stammfunktion für $u'v + uv'$ ist. Damit haben wir die folgende Regel:

Proposition 7.5.1 (partielle Integration). *Sind $u, v\colon I \to \mathbb{R}$ stetig differenzierbar, dann gilt*

$$\int uv' = uv - \int u'v, \qquad \int_a^b uv' = uv|_a^b - \int_a^b u'v.$$

Beispiel 7.5.2. *Für $a \neq -1$ gilt mit $u(x) = \ln(x)$ und $v(x) = \frac{1}{a+1}x^{a+1}$*

$$\int x^a \ln(x)\mathrm{d}x = \frac{1}{a+1}x^{a+1}\ln x - \frac{1}{a+1}\int x^a\mathrm{d}x = \frac{x^{a+1}}{(a+1)^2}\left((a+1)\ln(x) - 1\right).$$

Für $a = -1$ haben wir mit $u(x) = v(x) = \ln(x)$

$$\int \frac{\ln(x)}{x}\mathrm{d}x = \ln(x)^2 - \int\frac{\ln(x)}{x}\mathrm{d}x,$$

daraus folgt $\int \frac{\ln(x)}{x}\mathrm{d}x = \frac{\ln(x)^2}{2}$. □

Beispiel 7.5.3. *Wir wollen* $\int \cos^k x\mathrm{d}x$ *für natürliche Zahlen* $k \geq 2$ *berechnen. Mit partieller Integration mit* $u(x) = \cos^{k-1}(x)$ *und* $v(x) = \sin(x)$ *haben wir*

$$\begin{aligned}\int \cos^k x\mathrm{d}x &= \cos^{k-1}(x)\cdot\sin(x) + (k-1)\int \cos^{k-2}(x)\cdot\sin^2(x)\mathrm{d}x \\ &= \cos^{k-1}(x)\cdot\sin(x) + (k-1)\int (1-\cos^2(x))\cos^{k-2}(x)\mathrm{d}x.\end{aligned}$$

Umstellen liefert

$$\int \cos^k x\mathrm{d}x = \frac{1}{k}\cos^{k-1}(x)\cdot\sin(x) + \frac{k-1}{k}\int \cos^{k-2}(x)\mathrm{d}x.$$

Damit erhalten wir Rekursionsformeln, mit denen die unbestimmten Integrale ausgerechnet werden können.

Analog gilt

$$\int \sin^k x\mathrm{d}x = -\frac{1}{k}\sin^{k-1}(x)\cdot\cos(x) + \frac{k-1}{k}\int \sin^{k-2}(x)\mathrm{d}x.$$

Damit können zum Beispiel bestimmte Integrale $\int_0^{\frac{\pi}{2}} \cos^n(x)\mathrm{d}x$ *und* $\int_0^{\frac{\pi}{2}} \sin^n(x)\mathrm{d}x$ *berechnet werden.* □

7.5.2. Substitutionsregel.

Proposition 7.5.4. *Sei* $f\colon I \to \mathbb{R}$ *eine Regelfunktion mit Stammfunktion* F. *Sei* $g\colon [a,b] \to I$ *stetig differenzierbar und streng monoton. Dann ist* $F \circ g$ *eine Stammfunktion zu* $(f\circ g)\cdot g'$ *und es gilt*

$$\int_a^b f(g(x))\cdot g'(x)\mathrm{d}x = \int_{g(a)}^{g(b)} f(t)\mathrm{d}t.$$

Beispiel 7.5.5.

$$\int_a^b f(x+c)\mathrm{d}x = \int_{a+c}^{b+c} f(t)\mathrm{d}t.$$

Für $c \neq 0$ *gilt*

$$c\int_a^b f(cx)\mathrm{d}x = \int_{ac}^{bc} f(t)\mathrm{d}t.$$

□

Beispiel 7.5.6 (Logarithmische Integration).

$$\int \frac{g'(x)}{g(x)}\mathrm{d}x = \ln|g(x)|$$

□

Beispiel 7.5.7.

$$\int \frac{x}{x^2+1}\mathrm{d}x = \frac{1}{2}\ln(x^2+1).$$

□

Übungsaufgabe 7.5.8. *Sei* $f\colon [a,b] \to \mathbb{R}$ *eine stetig differenzierbare und streng monotone Funktion mit Stammfunktion* f. *Zeigen Sie die folgende Formel für die Stammfunktion* G *der Umkehrfunktion* f^{-1}*:*

$$G(y) = yf^{-1}(y) - F(f^{-1}(y)).$$

Hinweis: *Wenden Sie auf das Integral* $\int_a^b f^{-1}(y)\mathrm{d}y$ *die Substitution* $y = f(x)$ *an, danach partielle Integration.*

BEISPIEL 7.5.9.

$$\int \frac{Bx+C}{x^2+2bx+c}\mathrm{d}x$$

mittels Substitution □

7.5.3. Partialbruchzerlegung. Rationale Funktionen werden mittels Partialbruchzerlegung integriert. Nach Polynomdivision ist der relevante Fall

$$\int \frac{p(x)}{q(x)}\mathrm{d}x,$$

wobei p, q Polynome mit $\deg p < \deg q$ sind.

(1) Im ersten Schritt wird der Nenner in lineare und quadratische Faktoren zerlegt:

$$q(x) = a(x-b_1)^{m_1}\cdots(x-b_k)^{m_k}Q_1(x)^{n_1}\cdots Q_l(x)^{n_l}$$

wobei die b_i paarweise verschiedene Nullstellen mit Vielfachheiten m_i sind und die Q_j paarweise verschiedene quadratische Polynome, die keine reellen Nullstellen besitzen.

(2) Im zweiten Schritt setzen wir an

$$\frac{p(x)}{q(x)} = \sum_{i=1}^{k}\sum_{j=1}^{m_i}\frac{A_j}{(x-b_i)^j} + \sum_{i=1}^{l}\sum_{j=1}^{n_i}\frac{B_jx+C_j}{Q_i(x)^j}$$

Die Summe auf der rechten Seite wird auf den Hauptnenner $q(x)$ gebracht, und die Koeffizienten A_j, B_j, C_j werden durch Koeffizientenvergleich ermittelt.

(3) Danach können die einzelnen Summanden integriert werden. Die grundlegenden Integrale dafür sind

$$\begin{aligned}
\int \frac{\mathrm{d}x}{(x-a)^k} &= \frac{1}{(1-k)(x-a)^{k-1}} + C, k>1\\
\int \frac{\mathrm{d}x}{x-a} &= \ln|x-a| + C\\
\int \frac{\mathrm{d}x}{x^2+a^2} &= \frac{1}{a}\arctan\frac{x}{a} + C\\
\int \frac{bx+c}{x^2+a^2}\mathrm{d}x &= \frac{b}{2}\ln|x^2+a^2| + \frac{c}{a}\arctan\frac{x}{a} + C
\end{aligned}$$

Integrale für die allgemeinere Form x^2+px+q des Nenners können durch quadratische Ergänzung auf die Form $x^2+px+q = (x-b)^2+a^2$ gebracht und mit der Subsitution $y = x-b$ auf die obigen spezielleren Integrale zurückgeführt werden.

ÜBUNGSAUFGABE 7.5.10. *Sei $4q-p^2 > 0$, so dass das Polynom x^2+px+q keine reellen Nullstellen hat. Bestimmen Sie Rekursionsformeln zur Berechnung der Integrale*

$$\int \frac{\mathrm{d}x}{(x^2+a^2)^k}, \qquad \int \frac{bx+c}{(x^2+a^2)^k}\mathrm{d}x.$$

BEISPIEL 7.5.11. *Wir wollen die Funktion*

$$f(x) = \frac{x^3+4x^2}{(x^2+1)(x-5)^2}$$

integrieren. Hier haben wir bereits eine Faktorisierung des Nenners. Wir setzen an

$$\begin{aligned}\frac{x^3+4x^2}{(x^2+1)(x-5)^2} &= \frac{A_1}{x-5}+\frac{A_2}{(x-5)^2}+\frac{B_1x+C_1}{x^2+1}\\ &= \frac{A_1(x-5)(x^2+1)+A_2(x^2+1)+(B_1x+C_1)(x-5)^2}{(x^2+1)(x-5)^2}\end{aligned}$$

Nach Ausmultiplizieren ist der Zähler in der letzten Zeile

$(A_1+B_1)x^3+(-5A_1+A_2-10B_1+C_1)x^2+(A_1+25B_1-10C_1)x+(-5A_1+A_2+25C_1)$

und Koeffizientenvergleich mit x^3+4x^2 ergibt folgendes lineares Gleichungssystem:

$$\begin{aligned} A_1+B_1 &= 1\\ -5A_1+A_2-10B_1+C_1 &= 4\\ A_1+25B_1-10C_1 &= 0\\ -5A_1+A_2+25C_1 &= 0\end{aligned}$$

Mit dem Gauß-Algorithmus ergibt sich $A_1=\frac{185}{169}$, $A_2=\frac{225}{26}$, $B_1=-\frac{16}{169}$ und $C_1=-\frac{43}{338}$. Damit kann das Integral jetzt direkt hingeschrieben werden... □

Varianten über Exponentialfunktion und Winkelfunktionen mittels rationaler Parametrisierung der Kreislinie

Übungsaufgabe 7.5.12. *Berechnen Sie die folgenden unbestimmten Integrale mittels Partialbruchzerlegung:*

$$\int\frac{2x^2+2}{x^4-2x^2+1}\mathrm{d}x,\qquad \int\frac{5x^3-10x^2-13x-15}{x^4-x^3-x+1}\mathrm{d}x$$

7.5.4. Weitere Aufgaben.

Übungsaufgabe 7.5.13. *Berechnen Sie die folgenden bestimmten Integrale:*

$$\int_1^{\mathrm{e}^3}\frac{1}{x}\mathrm{d}x,\qquad \int_0^{2\pi}\sin^2(x)\mathrm{d}x,\qquad \int_1^{10}\sqrt{x-1}\cdot\ln(x-1)\mathrm{d}x,\qquad \int_0^{\frac{\pi}{2}}\sin(x)\cdot\cos(x)\mathrm{d}x.$$

$$\int_0^1 x^n(1-x)^m\mathrm{d}x,\qquad \int_{-1}^1(1+x)^n(1-x)^m\mathrm{d}x.$$

Übungsaufgabe 7.5.14. *Berechnen Sie die folgenden unbestimmten Integrale:*

$$\int\sqrt{x^2+4x+5}\mathrm{d}x,\qquad \int\sin(\sqrt[3]{x})\mathrm{d}x,\qquad \int\frac{\cos(x)}{1+\sin^2(x)}\mathrm{d}x,\qquad \int\frac{3}{2+\cos(x)}\mathrm{d}x.$$

Übungsaufgabe 7.5.15. *Berechnen Sie das Integral*

$$\int\frac{\mathrm{d}x}{ax^2+bx+c}$$

in Abhängigkeit von $a,b,c\in\mathbb{R}$,$a\neq 0$.

7.6. Uneigentliche Integrale

Bisher können wir nur Funktionen über abgeschlossene Intervalle integrieren. Durch geeignete Grenzwerte können auch manchmal allgemeinere Fälle behandelt werden:

Definition 7.6.1. *Sei f eine Regelfunktion auf einem Intervall I mit den Randpunkten a,b mit $-\infty\le a<b\le\infty$. Im Fall $I=[a,b)$ definieren wir*

$$\int_a^b f(x)\mathrm{d}x := \lim_{z\to b^-}\int_a^z f(x)\mathrm{d}x,$$

falls der Grenzwert auf der rechten Seite existiert. Wir sagen dann, dass das uneigentliche Integral $\int_a^b f(x)\mathrm{d}x$ konvergent *ist, der Grenzwert ist dann der* Wert *des uneigentlichen Integrals. Analog für den Fall* $(a, b]$. *Im Fall* $I = (a, b)$ *definieren wir*

$$\int_a^b f(x)\mathrm{d}x := \int_a^c f(x)\mathrm{d}x + \int_c^b f(x)\mathrm{d}x,$$

falls für ein c *die uneigentlichen Integrale auf der rechten Seite konvergent sind. Der Wert ist dann unabhängig von der Wahl von* c.

Beispiel 7.6.2. *Das uneigentliche Integral*

$$\int_1^\infty \frac{\mathrm{d}x}{x^s}$$

existiert genau dann, wenn $s > 1$. *Für* $s \neq 1$ *haben wir*

$$\int_1^t \frac{\mathrm{d}x}{x^s} = \frac{x^{1-s}}{1-s}\bigg|_1^t = \frac{t^{1-s}-1}{1-s}$$

Für $s > 1$ *ist der Grenzwert*

$$\lim_{t\to\infty} \frac{t^{1-s}-1}{1-s} = \frac{1}{s-1}.$$

Für $s < 1$ *existiert der Grenzwert nicht. Im Spezialfall* $s = 1$ *haben wir*

$$\int_1^t \frac{\mathrm{d}x}{x} = \ln(x)|_1^t = \ln(t),$$

und auch hier existiert der Grenzwert $\lim_{t\to\infty} \ln(t)$ *nicht.*

Umgekehrt existiert das uneigentliche Integral

$$\int_0^1 \frac{\mathrm{d}x}{x^s}$$

genau dann, wenn $s < 1$ *ist.* □

Proposition 7.6.3. *Seien* f *und* g *Regelfunktionen auf* $[a, b)$ *mit* $|f| \leq g$. *Wenn das Integral* $\int_a^b g(x)\mathrm{d}x$ *existiert, dann existiert auch* $\int_a^b f(x)\mathrm{d}x$.

Beispiel 7.6.4.

$$\int_0^\infty \frac{\sin(x)}{x}\mathrm{d}x.$$

□

Beispiel 7.6.5. *Gamma-Funktion*

$$\Gamma(x) := \int_0^\infty t^{x-1}\mathrm{e}^{-t}\mathrm{d}t$$

Diskussion Existenz

□

Übungsaufgabe 7.6.6. *Funktionalgleichung für Gamma-Funktion, Identifikation der Funktionswerte mit Fakultät*
Hinweis: *partielle Integration*

Übungsaufgabe 7.6.7. *Fehlerintegral über Gamma-Funktion* Hinweis: *Substitution*

Übungsaufgabe 7.6.8. *(1) Berechnen Sie die folgenden uneigentlichen Integrale, falls sie existieren:*

$$\int_0^{+\infty} x^2 \mathrm{e}^{-x} \mathrm{d}x, \qquad \int_1^{+\infty} \frac{2}{x^2+4x+3} \mathrm{d}x$$

$$\int_0^{+\infty} \mathrm{e}^{-ax} \mathrm{d}x \text{ für } a \in \mathbb{R}, \qquad \int_{-1}^{1} \frac{-2x}{1-x^2} \mathrm{d}x, \qquad \int_{-\infty}^{+\infty} x \mathrm{e}^{x^2} \mathrm{d}.$$

(2) Für welche $\alpha \in \mathbb{R}$ existiert das uneigentliche Integral $\int_0^{+\infty} x^\alpha \mathrm{d}x$?

Übungsaufgabe 7.6.9. *Berechnen Sie das uneigentliche Integral*

$$\int_0^1 \ln(x)^n \mathrm{d}x.$$

7.7. Anwendungen

Bemerkungen zu komplexen Funktionen
Analysis 2: Volumina, Oberflächen, Bogenlängen

7.7.1. Flächenberechnung.

Beispiel 7.7.1. *Der Einheitskreis ist die von $x^2 + y^2 = 1$ begrenzte Fläche in der Ebene. Die obere Hälfte des Einheitskreises wird also vom Graphen der Funktion $y = \sqrt{1-x^2}$ begrenzt. Die Fläche des Einheitskreises ist also gleich*

$$2 \int_{-1}^{1} \sqrt{1-x^2} \mathrm{d}x$$

Für dieses Integral benutzen wir die Substitution $x = \cos t$ (das bedeutet, dass wir die x-Koordinate durch den Winkel t ausdrücken). Dann ist $\mathrm{d}x = -\sin(t)\mathrm{d}t$, und wir rechnen

$$\begin{aligned}
\int_{-1}^{1} \sqrt{1-x^2} \mathrm{d}x &= \int_{\arccos(-1)}^{\arccos(1)} \sqrt{1-\cos^2 t}(-\sin t)\mathrm{d}t \\
&= \int_{\pi}^{0} \sin t \cdot (-\sin t)\mathrm{d}t \\
&= \int_0^{\pi} \sin^2 t \mathrm{d}t \\
&= \frac{t - \sin t \cdot \cos t}{2}\Big|_0^{\pi} = \frac{\pi}{2}.
\end{aligned}$$

Dabei benutzen wir im letzten Schritt die Rekursionsformel aus Beispiel 7.5.3. Der Flächeninhalt des Einheitskreises ist also π. □

Übungsaufgabe 7.7.2. *Seien $a, b > 0$ reelle Zahlen. Berechnen Sie den Flächeninhalt der Ellipse*

$$\left\{ (x,y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1 \right\}$$

.

7.7.2. Irrationalität von π.

7.7.3. Ableitung und Integral für Funktionenfolgen.

PROPOSITION 7.7.3. *Sei $(f_n)_n$ eine gleichmäßig konvergente Folge von Regelfunktionen $f_n\colon [a,b] \to \mathbb{R}$. Dann ist $f = \lim_{n\to\infty} f_n$ eine Regelfunktion auf $[a,b]$ und es gilt*

$$\int_a^b f(x)\mathrm{d}x = \lim_{n\to\infty} \int_a^b f_n(x)\mathrm{d}x.$$

BEWEIS. Nach Voraussetzung der gleichmäßigen Konvergenz existiert für jedes $\epsilon > 0$ ein n mit $\|f - f_n\|_\infty < \frac{\epsilon}{2}$. Da f_n eine Regelfunktion ist, existiert auch eine Treppenfunktion ϕ mit $\|f_n - \phi\|_\infty < \frac{\epsilon}{2}$. Dann ist $\|f - \phi\| < \epsilon$ und wir sehen, dass f eine Regelfunktion ist. Die Behauptung folgt nun aus der Beschränktheit des Regelintegrals:

$$\left| \int_a^b f(x)\mathrm{d}x - \int_a^b f_n(x)\mathrm{d}x \right| \leq \|f - f_n\|_\infty (b-a).$$

□

BEISPIEL 7.7.4. *Für eine Potenzreihe*

$$f(x) = \sum_{n=0}^\infty a_n x^n$$

mit Konvergenzradius R gilt für jedes $x \in (-R, R)$

$$\int_0^x f(t)\mathrm{d}t = \sum_{n=0}^\infty \frac{a_n}{n+1} x^{n+1}.$$

□

BEISPIEL 7.7.5. *Für das* GAUSS*sche Fehlerintegral ergibt sich damit die Potenzreihenentwicklung*

$$\int_0^x \mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t = \sum_{n=0}^\infty \frac{(-1)^n x^{2n+1}}{2^n n!(2n+1)}.$$

□

BEISPIEL 7.7.6.

$$\mathrm{Li}_2(z) = -\int_0^z \frac{\ln(1-u)}{u}\mathrm{d}u = \sum_{k=1}^\infty \frac{z^k}{k^2}.$$

□

PROPOSITION 7.7.7. *Sei $(f_n\colon I \to \mathbb{R})_n$ eine Folge von stetig differenzierbaren Funktionen auf einem abgeschlossenen Intervall I, so dass die beiden folgenden Aussagen gelten:*

(1) Die Folge f_n konvergiert punktweise gegen die Funktion $f\colon [a,b] \to \mathbb{R}$.
(2) Die Folge der Ableitungen $f_n'\colon [a,b] \to \mathbb{R}$ konvergiert gleichmäßig.

Dann ist $f = \lim_{n\to\infty} f_n$ stetig differenzierbar und es gilt

$$f'(x) = \lim_{n\to\infty} f_n'(x)$$

BEWEIS. Wir setzen $g = \lim_{n\to\infty} f'$. Nach Proposition 4.3.4 ist g als gleichmäßiger Grenzwert stetiger Funktionen stetig. Für festes $z \in [a,b]$ und beliebiges $x \in [a,b]$ nach dem Hauptsatz der Differential- und Integralrechnung

$$f_n(x) = f_n(z) + \int_z^x f_n'(t)\mathrm{d}t.$$

Nach Proposition 7.7.3 folgt dann für $n \to \infty$

$$f(x) = f(z) + \int_z^x g(t)\mathrm{d}t$$

Nach dem Hauptsatz der Differential- und Integralrechnung ist dann f differenzierbar mit Ableitung $f'(x) = \lim_{n\to\infty} f'_n(x)$. □

Beispiel 7.7.8. *Ein Beispiel, in dem Proposition 7.7.7 nicht angewendet werden kann ist die Funktionenfolge* $f_n = \frac{1}{n}\sin(nx)$. *Diese Funktionenfolge konvergiert gleichmäßig gegen* 0, *aber die Folge* $f'_n = \cos(nx)$ *der Ableitungen konvergiert nicht. Es reicht also nicht aus, nur gleichmäßige Konvergenz für eine Funktionenfolge zu fordern, um Grenzwert und Ableitung vertauschen zu können.* □

Bemerkung 7.7.9. *Für* $0 < a < 1$ *und* $b > 0$ *mit* $ab \geq 1$ *betrachten wir die Weierstraß-Funktion*

$$f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x).$$

Die Weierstraß-Funktion ist auf ganz $\mathbb{R}$ *stetig, aber nirgends differenzierbar. Stetigkeit ist eine einfache Konsequenz aus der gleichmäßigen Konvergenz. Dass* f *nirgends differenzierbar ist, wurde von Hardy bewiesen.*

Übungsaufgabe 7.7.10.

$$\int_0^1 x^{-x}\mathrm{d}x = \sum_{n=1}^{\infty} n^{-n}$$

KAPITEL 8

Anwendungen von Potenzreihen

8.1. Taylor-Entwicklung

Sei $f\colon D \to \mathbb{R}$ eine im Punkt $c \in D$ n-mal differenzierbare Funktion. Wir wollen wissen, ob bzw. wann sich f durch einfachere Funktionen (Polynome) annähern lässt. Die beste lineare Annäherung hatten wir bereits in Bemerkung 6.1.5 gesehen, gegeben durch die Tangente mit der Gleichung

$$y = f(c) + f'(c)(x - c)$$

Allgemeiner suchen wir jetzt für die n-mal differenzierbare Funktion f ein Polynom $T(x)$ vom Grad n mit

$$T(c) = f(c), T'(c) = f'(c), \qquad T^{(n)}(c) = f^{(n)}(c).$$

Wir machen den Ansatz $f(x) = T(x) := \sum_{j=0}^{n} a_j(x - c)^j$ und erhalten wir durch Ableiten der Gleichung

$$f^{(k)}(x) = \sum_{j=k}^{n} j(j-1)(j-2)\cdots(j-k+1)a_j(x-c)^{j-k}.$$

Für $x = c$ gilt dann insbesondere $f^{(k)}(c) = k!a_k$. Das bedeutet, dass ein Polynom vom Grad n, dessen erste Ableitungen am Punkt c mit den Ableitungen von f übereinstimmen eindeutig bestimmt ist und die folgende Form hat:

$$T_{f,n,c}(x) = f(c) + \frac{f'(c)}{1!}(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots \frac{f^{(n)}(c)}{n!}(x-c)^n.$$

Das Polynom $T_{f,n,c}(x)$ heißt n-tes *Taylorpolynom* für die Funktion f im Punkt c. Die Polynome $T_{f,n,c}$ schmiegen sich in einer Umgebung des Punktes c mit wachsendem n immer besser an den Graph der Funktion f an, da die Werte für immer höhere Ableitungen übereinstimmen, s. Abbildungen 27 und 28. Für $n = 1$ haben wir den richtigen Anstieg, für $n = 2$ die richtige Krümmung/Konvexität,...

Satz 8.1.1. *Für jede $(n+1)$-mal stetig differenzierbare Funktion $f\colon I \to \mathbb{R}$ auf einem offenen Intervall $I \subseteq R$ und $c \in I$ gilt*

$$f(x) = \underbrace{\sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k}_{\textit{Taylorpolynom}} + \underbrace{R_{n+1}(x,c)}_{\textit{Restglied}}$$

Integraldarstellung Restglied:

$$R_{n+1}(x,c) = \frac{1}{n!}\int_c^x (x-t)^n f^{(n+1)}\mathrm{d}t$$

Lagrangedarstellung Restglied: *Für ein geeignetes $\xi \in [c, x]$ gilt*

$$R_{n+1}(x,c) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1}$$

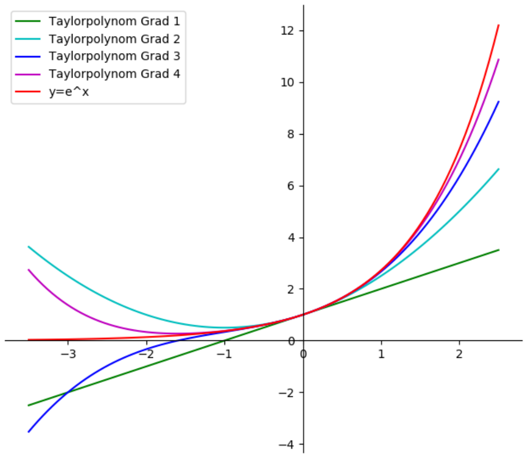


ABBILDUNG 27. Taylorpolynome von Grad 1 bis 4 für Exponentialfunktion $\exp(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \cdots$ um den Entwicklungspunkt 0

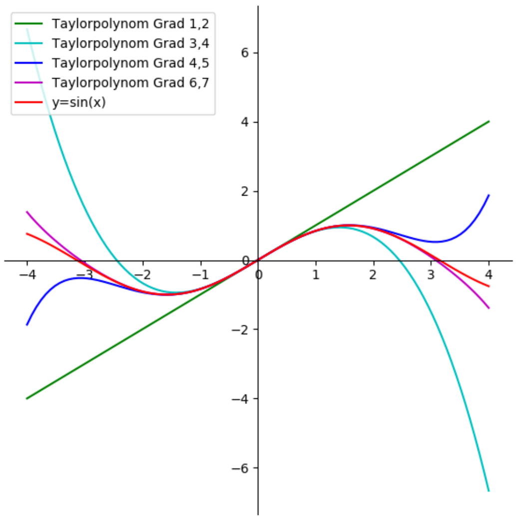


ABBILDUNG 28. Taylorpolynome von Grad 1 bis 7 für die Sinusfunktion $\sin(x) = x - \frac{x^3}{6} + \frac{x^5}{120} - \frac{x^7}{5040} + \cdots$ um den Entwicklungspunkt 0

BEWEIS. Für die Integraldarstellung des Restglieds benutzen wir eine Induktion über n. Der Fall $n = 0$ folgt direkt aus dem Hauptsatz der Differential- und Integralrechnung:

$$f(x) = f(c) + \int_c^x f'(t)\mathrm{d}t$$

Im Induktionsschritt setzen wir die Behauptung für $n-1$ voraus:

$$f(x) - T_{n-1}(x) = \frac{1}{(n-1)!}\int_c^x (x-t)^{n-1} f^{(n)}(t)\mathrm{d}t.$$

Mit $u = -(x-t)^n$ (Integrationsvariable t!)und $v = f^{(n)}$ erhalten wir durch partielle Integration

$$f(x) - T_{n-1}(x) = \frac{f^{(n)}(c)}{n!}(x-c)^n + \frac{1}{n!}\int_c^x (x-t)^n f^{(n+1)}(t)\mathrm{d}t.$$

Die Lagrangeform folgt aus dem Mittelwertsatz der Integralrechnung

$$\frac{1}{n!}\int_c^x (x-t)^n f^{(n+1)}(t)\mathrm{d}t = \frac{1}{n!}f^{(n+1)}(\xi)\int_c^x (x-t)^n \mathrm{d}t = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1},$$

wobei $\xi \in [c, x]$ □

KOROLLAR 8.1.2. *Eine Funktion $f : I \to \mathbb{R}$, die $n+1$-mal differenzierbar ist, und für die gilt $f^{(n+1)}(x) = 0$, ist ein Polynom.*

Qualitative Form der Taylorformel:

KOROLLAR 8.1.3. *Sei $f : I \to \mathbb{R}$ eine n-mal stetig differenzierbare Funktion und $c \in I$. Dann gilt für alle $x \in I$*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k + \phi(x)(x-c)^n$$

wobei ϕ eine stetige Funktion mit $\lim_{x\to c}\phi(x) = 0$ ist.

Umformuliert, sagt diese Formel, dass f mit dem n-ten Taylorpolynom in n-ter Näherung übereinstimmt. Der Unterschied zwischen f und T_n ist von der Größenordnung $|x-c|^n$.

BEISPIEL 8.1.4. *Die qualitative Taylorformel für die Wurzelfunktion $f : (-1,1) \to \mathbb{R} : x \mapsto \sqrt{1+x}$ um den Entwicklungspunkt $c = 0$ liefert*

$$f(0) = 1, f'(0) = \frac{1}{2\sqrt{1+x}}\Big|_{x=0} = \frac{1}{2},$$

und damit $\sqrt{1+x} = 1 + \frac{x}{2} + \phi(x)x$ für eine stetige Funktion $\phi(x)$ mit $\lim_{x\to 0}\phi(x) = 0$. □

ÜBUNGSAUFGABE 8.1.5. *Benutzen Sie die qualitative Taylorformel für $f(x) = \sqrt{1+x}$, um die folgende Aussage zu zeigen:*

$$\lim_{n\to\infty}\left(\sqrt{n+\sqrt{n}} - \sqrt{n}\right) = \frac{1}{2}$$

Um eine Funktion immer besser annähern zu können gehen wir von Polynomen zu Potenzreihen über und fragen uns, unter welchen Bedingungen eine Funktion in einer Umgebung eines Punktes als Potenzreihe geschrieben werden kann.

DEFINITION 8.1.6. *Eine unendliche Reihe der Form $\sum_{k=0}^{\infty} a_k(x-c)^k$ heißt Potenzreihe mit Entwicklungspunkt c.*

Der Entwicklungspunkt der Potenzreihe kann unter Konvergenzvoraussetzungen gewechselt werden:

PROPOSITION 8.1.7. *Sei $\sum_{n=0}^{\infty} a_n(x-c)^n$ eine Potenzreihe mit Konvergenzradius R und d ein Punkt mit $|d-c| < R$. Dann gilt für alle x mit $|x-d| < R - |d-c|$ die Gleichung*

$$\sum_{n=0}^{\infty} a_n(x-c)^n = \sum_{k=0}^{\infty} b_k(x-d)^k, \qquad \textit{mit} \qquad b_k = \sum_{k=n}^{\infty}\binom{n}{k} a_n(d-c)^{n-k}.$$

Beweis. Im Konvergenzkreis haben wir durch Ausmultiplizieren

$$\sum_{n=0}^{\infty} a_n(x-c)^n = \sum_{n=0}^{\infty} a_n\left((x-d)+(d-c)\right)^n = \sum_{n=0}^{\infty}\sum_{k=0}^{n} a_n \binom{n}{k}(d-c)^{n-k}(x-d)^k.$$

Durch die absolute Konvergenz im Inneren des Konvergenzkreises können wir die Summationsreihenfolge ändern (dabei wird die $\binom{n}{k} = 0$ für $k > n$ benutzt):

$$\begin{aligned}\sum_{n=0}^{\infty}\sum_{k=0}^{n} a_n \binom{n}{k}(d-c)^{n-k}(x-d)^k &= \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} a_n \binom{n}{k}(d-c)^{n-k}(x-d)^k \\ &= \sum_{k=0}^{\infty}\left(\sum_{n=0}^{\infty} a_n \binom{n}{k}(d-c)^{n-k}\right)(x-d)^k\end{aligned}$$

□

Beispiel 8.1.8. *Für die Exponentialfunktion benutzen wir die Funktionalgleichung* $\mathrm{e}^x = \mathrm{e}^c\mathrm{e}^{x-c}$, *um den Entwicklungspunkt zu ändern:*

$$\mathrm{e}^x = \sum_{k=0}^{\infty} \frac{\mathrm{e}^c}{k!}(x-c)^k.$$

□

Bemerkung 8.1.9. *Durch Umparametrisieren der Potenzreihe zu einem anderen Entwicklungspunkt kann sich der Konvergenzradius ändern.*

Übungsaufgabe 8.1.10. *Bestimmen Sie den Konvergenzradius der Potenzreihe, die sich aus der geometrischen Reihe durch Umparametrisierung mit dem Entwicklungspunkt* $x = -\frac{1}{2}$ *ergibt.*

Hinweis: *Beweisen Sie dafür die Formel*

$$\frac{1}{(1-x)^{k+1}} = \sum_{n=0}^{\infty} \binom{n+k}{n} x^n$$

für $k \in \mathbb{N}$ *und* $|x| < 1$, *und benutzen Sie diese.*

Aus dem Identitätssatz für Potenzreihen, cf. Korollar 3.2.13, folgt:

Proposition 8.1.11. *Wenn eine Funktion* $f\colon (c-R, c+R) \to \mathbb{R}$ *als Potenzreihe mit Entwicklungspunkt* c *und Konvergenzradius* $\geq R$ *geschrieben werden kann, dann ist die Darstellung als Potenzreihe eindeutig:*

$$f(x) = \sum_{k=0}^{\infty} a_k(x-c)^k, \qquad a_k = \frac{f^{(k)}(c)}{k!}.$$

Definition 8.1.12. *Sei* $f\colon I \to \mathbb{R}$ *eine beliebig oft differenzierbare Funktion und* $c \in I$. *Die Potenzreihe*

$$T_f(x,c) = \sum_{k=0}^{\infty} \frac{f^{(k)}(c)}{k!}(x-c)^k$$

heißt Taylorreihe zu f mit Entwicklungspunkt c.

Wenn $T_f(x,c)$ *Konvergenzradius* $R > 0$ *hat und für alle* $x \in (c-R, c+R) \cap I$ *gilt* $f(x) = T_f(x,c)$, *dann sagen wir, dass sich* f *um* c *als Taylorreihe entwicklen lässt.*

Bemerkung 8.1.13. *Zu beachten ist:*

(1) *Der Konvergenzradius der Taylorreihe ist nicht notwendig* > 0.
(2) *Auch wenn die Taylorreihe zu* f *einen positiven Konvergenzradius hat, konvergiert sie nicht notwendig gegen* f.

Funktionen, die sich in einer Umgebung eines Punktes in eine Potenzreihe entwickeln lassen, heißen auch analytische Funktionen*.*

ÜBUNGSAUFGABE 8.1.14. *Bestimmen Sie die Taylorreihe zu* $f(x) = \mathrm{e}^{-\frac{1}{x^2}}$ *mit Entwicklungspunkt* $c = 0$*. Lässt sich* f *um* 0 *als Taylorreihe entwickeln?*

ÜBUNGSAUFGABE 8.1.15. *Bestimmen Sie die Taylorentwicklung der Funktion* $f(x) = \sin^2(x)$ *um den Entwicklungspunkt* 0*. (Varianten: Taylorentwicklung für Komposition durch Einsetzen, oder Bestimmung der Taylorreihe für die Ableitung, dann Integration.)*

SATZ 8.1.16. *Sei* $f\colon I \to \mathbb{R}$ *beliebig oft differenzierbar und* $c \in I$*. Die Taylorreihe* $T_f(x,c)$ *konvergiert genau dann gegen* $f(x)$ *für* $x \in I$*, wenn die Folge* $(R_n(x,c))_n$ *der Restglieder eine Nullfolge ist. Eine hinreichende Bedingung dafür ist, dass* A, B *existiert, so dass für alle* $n \in \mathbb{N}$ *und alle* $x \in I$ *gilt*

$$|f^{(n)}(x)| \leq AB^n.$$

BEWEIS. Die Äquivalenz folgt aus $f(x) - T_{f,n,c}(x) = R_{n+1}(x,c)$. Aus $|f^{(n)}(x)| \leq AB^n$ folgt

$$|R_n(x,c)| = \left| \frac{f^{(n)}(\xi)}{n!}(x-c)^n \right| \leq \frac{AB^n}{n!}|x-c|^n.$$

Aus der Konvergenz der Exponentialreihe

$$\sum_{k=0}^{\infty} A \frac{B^n |x-c|^n}{n!}$$

(gegen $A\mathrm{e}^{B|x-c|}$) sehen wir, dass $\frac{AB^n|x-c|^n}{n!}$ eine Nullfolge ist. $\square$

BEISPIEL 8.1.17. *Sei* $|x| < 1$ *und* $\alpha \in \mathbb{R}$*. Wir bestimmen die Taylorentwicklung der Funktion* $f(x) = (1+x)^\alpha$*. Dafür setzen wir*

$$\binom{\alpha}{k} := \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}.$$

Dies verallgemeinert die Binomialkoeffizienten (im Spezialfall $\alpha \in \mathbb{N}$*). Eine explizite Berechnung der Ableitungen von* $f(x)$ *zeigt*

$$T_f(x,0) = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k.$$

Diese Reihe heißt auch Binomialreihe*.*

Wir bestimmen den Konvergenzradius mit Hilfe des Quotientenkriteriums:

$$\left| \frac{a_{k+1}}{a_k} \right| = \frac{\alpha(\alpha-1)\cdots(\alpha-(k+1)+1)k!}{\alpha(\alpha-1)\cdots(\alpha-k+1)(k+1)!} = \frac{\alpha-k}{k+1}.$$

Dann ist $\lim_{k\to\infty} \frac{\alpha-k}{k+1} = -1$*, und damit ist der Konvergenzradius der Binomialreihe* $R = 1$*.*

Wir müssen noch zeigen, dass die Binomialreihe tatsächlich gegen die Funktion $f(x) = (1+x)^\alpha$ *konvergiert. Statt einer Restgliedbetrachtung setzen wir* $g(x) = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k$*. Dann ist* $g'(x) = \sum_{k=1}^{\infty} k\binom{\alpha}{k} x^{k-1}$*, woraus* $g'(x)(1+x) = \alpha g(x)$ *folgt. Mit der Quotientenregel ist*

$$\begin{aligned} \left(\frac{g(x)}{(1+x)^\alpha} \right)' &= \frac{g'(x)(1+x)^\alpha - g(x)\alpha(1+x)^{\alpha-1}}{(1+x)^{2\alpha}} \\ &= \frac{(1+x)^{\alpha-1}\left(g'(x)(1+x) - \alpha g(x)\right)}{(1+x)^{2\alpha}} = 0 \end{aligned}$$

Damit ist $\frac{g(x)}{(1+x)^\alpha}$ *konstant. Wegen* $f(0) = g(0) = 1$ *folgt dann* $g(x) = (1+x)^\alpha$*.*

Die binomische Formel für $\alpha \in \mathbb{N}$ ist dann ein Spezialfall der binomischen Reihe. Als weiteren Spezialfall erhalten wir insbesondere Taylor-Entwicklungen für die Wurzelfunktionen.

- *Für $\alpha = \frac{1}{2}$ gilt $\binom{\frac{1}{2}}{0} = 1$ und $\binom{\frac{1}{2}}{1} = \frac{1}{2}$, und*

$$\binom{\frac{1}{2}}{k} = (-1)^{k-1} \frac{1 \cdot 3 \cdot 5 \cdots (2k-3)}{2 \cdot 4 \cdots (2k)}.$$

Damit haben wir für $|x| < 1$

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4 + \cdots .$$

- *Für $\alpha = -\frac{1}{2}$ gilt $\binom{-\frac{1}{2}}{0} = 1$ und*

$$\binom{-\frac{1}{2}}{k} = (-1)^{k} \frac{1 \cdot 3 \cdot 5 \cdots (2k-1)}{2 \cdot 4 \cdots (2k)}.$$

Damit haben wir für $|x| < 1$

$$\frac{1}{\sqrt{1+x}} = 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{5}{16}x^3 + \cdots$$

□

ÜBUNGSAUFGABE 8.1.18. *(1) Bestimmen Sie die Taylor-Entwicklung für die Funktion*

$$c(x) = \frac{1 \pm \sqrt{1-4x}}{2x}.$$

(2) Zeigen Sie

$$\binom{\frac{1}{2}}{n} = \frac{(-1)^{n+1}}{4^n(2n-1)}\binom{2n}{n}$$

(3) Leiten Sie daraus die explizite Formel für die Catalan-Zahlen ab, s. auch Übungsaufgabe 3.2.16:

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

Bernoulli-Zahlen und Cotangens-Reihe

ÜBUNGSAUFGABE 8.1.19. *Bestimmen Sie das Taylorpolynom $T_3(x, 0)$ vom Grad 3 mit Entwicklungspunkt $c = 0$ für die Funktion $f(x) = \frac{x}{\mathrm{e}^x - 1}$.*

8.2. Taylor-Entwicklung für Näherungsrechnungen

BEISPIEL 8.2.1. *Wir wollen Werte der Exponentialfunktion e^x für $|x| \le 1$ mit Genauigkeit 10^{-3} berechnen (also die ersten beiden Nachkommastellen). Aus der Lagrange-Restgliedformel haben wir*

$$\mathrm{e}^x = \sum_{k=0}^{n} \frac{x^k}{k!} + \frac{\mathrm{e}^\xi}{(n+1)!}x^{n+1},$$

und wir benötigen damit eine Abschätzung für den letzten Term. Für $|x| \le 1$ ist $|\mathrm{e}^x| < 3$ und $|x^{n+1}| \le 1$, also

$$\frac{|\mathrm{e}^\xi|}{(n+1)!}|x^{n+1}| < \frac{3}{(n+1)!}.$$

Mit $\frac{3}{7!} = \frac{1}{1680} < \frac{1}{1000}$ *reicht also das Taylorpoylynom vom Grad 6 aus, um* e^x *für* $|x| \leq 1$ *auf 2 Nachkommastellen genau zu berechnen. Um die ersten 80 Nachkommastellen von* e *zu berechnen reicht wegen* $60! \sim 8.3 \cdot 10^{81} \geq 3 \cdot 10^{81}$ *das Taylorpolynom vom Grad 59. Die Exponentialreihe konvergiert damit sehr schnell und ist gut geeignet für numerische Näherungsrechnungen.* □

Mit dieser Abschätzung kann gezeigt werden, dass e irrational ist.

ÜBUNGSAUFGABE 8.2.2. *Berechnen Sie* $\sqrt{\mathrm{e}}$ *mit einer Genauigkeit von* 10^{-3}.
Hinweis: *Lagrange-Restglied der Taylorreihe.*

BEISPIEL 8.2.3. *Wir benutzen die Taylorentwicklung von* $f(x) = \ln(1+x)$*, um* $\ln(2)$ *näherungsweise zu berechnen. Mit vollständiger Induktion können wir sehen, dass*

$$f^{(k)}(x) = (-1)^{k-1} \frac{(k-1)!}{(1+x)^k}$$

ist. Daraus ergibt sich als Taylorentwicklung mit Entwicklungspunkt $x = 0$

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}(k-1)!}{k!} x^k = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k.$$

Dies ist die Logarithmus-Reihe, s.a. Übungsaufgabe 6.4.4. Der Konvergenzradius ist 1*. Nach der Lagrange-Restgliedformel ist*

$$\ln(1+x) = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k} x^k + \frac{(-1)^n}{n+1} \cdot \left(\frac{x}{\xi + 1} \right)^{n+1}$$

Für $\xi \in [0,1]$ *ist*

$$\left| \frac{(-1)^n}{n+1} \cdot \left(\frac{1}{\xi+1} \right)^{n+1} \right| \leq \frac{1}{n+1}.$$

Um $\ln(2)$ *mit einer Genauigkeit von* 10^{-3} *(also 2 Nachkommastellen) zu berechnen, reicht das Taylorpolynom vom Grad* n *mit*

$$\frac{1}{n+1} \leq 10^{-3},$$

also $n \geq 999$*. Die Logarithmusreihe konvergiert also eher langsam, nicht so geeignet für numerische Rechnungen. Allerdings ist auch die Abschätzung aus der Taylorformel nicht optimal, tatsächlich stimmen die ersten beiden Nachkommastellen bereits ab* $n = 160$*, die dritte Nachkommastelle ist dann ab* $n = 3398$ *richtig.*

Eine bessere Reihe für die Berechnung von $\ln(2)$ *ist*

$$\ln\left(\frac{1+x}{1-x}\right) = 2 \sum_{k=0}^{\infty} \frac{x^{2k+1}}{2k+1}$$

mit $x = \frac{1}{3}$. □

ÜBUNGSAUFGABE 8.2.4. *Berechnen Sie mit der normalen Logarithmusreihe eine Näherung von* $\ln(2)$ *als Bruch mit einer Genauigkeit von* 0.2.

Berechnen Sie $\ln(2)$ *mit der Taylorreihe für* $\ln\left(\frac{1+x}{1-x}\right)$ *bis auf zwei Nachkommastellen genau.*

BEISPIEL 8.2.5. *Eine praktisch sehr relevante und einfache Anwendung der Taylor-Reihe für den Logarithmus ist die folgende Abschätzung. Wir nehmen einen exponentiellen Wachstumsprozess mit Wachstumsrate* r*; ein Zuwachs von* a *Prozent pro Wachstumszyklus entspricht dann* $r = \frac{a}{100}$*. Wir wollen wissen, wie viele*

Wachstumszyklen bis zur Verdoppelung benötigt werden, wir suchen also n in der Gleichung $(1+r)^n x = 2x$. Die Lösung ist unabhängig von der Ausgangsgröße x:

$$n = \frac{\ln 2}{\ln(1+r)}.$$

Mit der Taylorreihe für den Logarithmus können wir also bei kleinen Wachstumsraten $\ln(1+r)$ durch r ersetzen. Benutzen wir $\ln(2) \approx 0.69$, lautet die Faustregel also: bei hinreichend kleinen Wachstumsraten von a Prozent verdoppelt sich die Menge nach $\frac{69}{a}$ Wachstumszyklen. *Man kann damit ohne Taschenrechner und große Überlegungen beispielsweise folgende einfache Abschätzungen vornehmen:*

- *Corona-Pandemie: Wächst die Anzahl der Infizierten pro Tag um 5%, verdoppelt sich die Zahl der Infizierten in 14 Tagen.*
- *Subprime-Krise: Nimmt man einen Kredit zu 2% Zinsen p.a. und einer Laufzeit von 35 Jahren auf, bezahlt man ein Haus für sich selbst und eins für den Bankberater.*
- *Nachhaltigkeit: Bei einem Wirtschaftswachstum von 2% pro Jahr verdoppelt sich alle 35 Jahre die Wirtschaftsleistung und damit gekoppelt der Ressourcenverbrauch; es ist ziemlich einfach zu sehen, dass das kein nachhaltiges Modell ist.*

Wir diskutieren noch den Fehler, der bei der Abschätzung $\ln(1+x) \approx x$ gemacht wird. Nach der Lagrange-Restgliedformel ist

$$\ln(1+x) = x - \frac{1}{2} \cdot \left(\frac{x}{\xi+1}\right)^2$$

für ein $\xi \in [0, x]$. Für $|x| \le 0.1$ ist dann

$$\frac{1}{2}\left(\frac{x}{\xi+1}\right)^2 \le 0.01.$$

Betrachten wir, wie sich ein solcher Fehler in der Näherungsformel auswirkt, sehen wir, dass für Wachstumsraten bis 10% der Fehler der Abschätzung weniger als ein Wachstumszyklus beträgt:

$$\frac{69}{10} = 6.9, \frac{69}{9} \approx 7.66, \frac{69}{11} \approx 6.27$$

Für eine solche Betrachtung kann auch wieder eine Taylorentwicklung für die Funktion $g(x) = \frac{\ln(2)}{x}$ mit Entwicklungspunkt $x = 0.1$ gemacht werden. □

ÜBUNGSAUFGABE 8.2.6. *näherungsweise Berechnung von $\sqrt{2}$*

ÜBUNGSAUFGABE 8.2.7. *Schätzen Sie ab, für welche Winkel ϕ bei der näherungsweisen Berechnung*

$$\sin\phi \approx \phi - \frac{\phi^3}{6} + \frac{\phi^5}{120}$$

der Fehler unter 10^{-6} liegt!

Hinweis: *Interpretieren Sie den Ausdruck als Taylorpolynom möglichst hohen Grades.*

ÜBUNGSAUFGABE 8.2.8.
(1) *Bestimmen Sie das Taylorpolynom $T_5(x, \pi)$ vom Grad 5 mit Entwicklungspunkt $c = \pi$ für die Funktion $f(x) = \tan(x)$.*
(2) *Berechnen Sie mit Hilfe von (2) näherungsweise das Integral*

$$\int_0^{\frac{\pi}{4}} \tan(x)\mathrm{d}x$$

Runden Sie dabei auf drei Dezimalstellen nach dem Komma. Vergleichen Sie das Ergebnis mit dem exakten Wert.

8.3. Komplexbeispiel: Periode des Fadenpendels

Herleitung der Bewegungsgleichung, Differentialgleichungen. Wir betrachten ein Fadenpendel: ein Objekt der Masse m ist an einem Faden der Länge l aufgehängt, und wird um den Winkel ϕ ausgelenkt. Um die Bewegung des Fadenpendels herzuleiten, betrachten wir die Kräfte, die auf die Masse m wirken; dabei werden Energieverluste durch Reibung vernachlässigt. Auf die Masse m wirkt die Schwerkraft $F_{\text{grav}} = mg$, wobei $g \approx 9.81 \frac{m}{s^2}$ die Fallbeschleunigung ist. Wesentlich für die Bewegung des Fadenpendels ist die Rückstellkraft, die senkrecht zum Faden in Bewegungsrichtung zeigt. Schwerkraft und Rückstellkraft bilden Hypotenuse bzw. Kathete eines rechtwinkligen Dreiecks, damit ist die Rückstellkraft bei einer Auslenkung um den Winkel ϕ gleich $F = mg\sin(\phi)$.

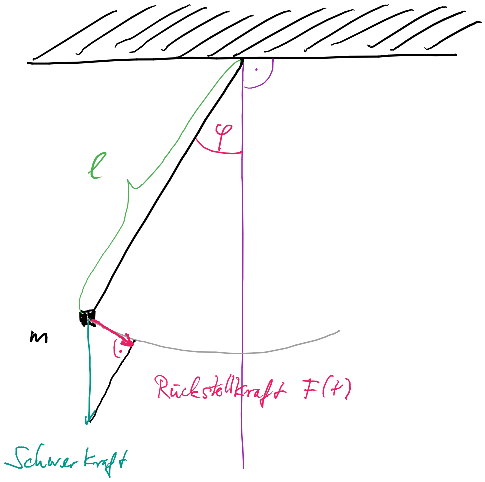


Abbildung 29. Kräfte am Fadenpendel

Für die Bewegungsgleichung wollen wir den Auslenkungswinkel $\phi(t)$ in Abhängigkeit von der Zeit beschreiben, die Rückstellkraft in Abhängigkeit von der Zeit ist dann

$$F(t) = -mg\sin(\phi(t)).$$

Das Vorzeichen besagt dabei, dass die Rückstellkraft in Richtung Ruhelage zeigt. Die Rückstellkraft führt zu einer Beschleunigung $a(t)$, und nach dem zweiten Newtonschen Gesetz gilt $F(t) = ma(t)$. Wir beschreiben die Beschleunigung als Winkelbeschleunigung, d.h.

$$a(t) = l\frac{\mathrm{d}^2\phi}{\mathrm{d}t^2}.$$

Kombinieren wir die beiden Gleichungen für Kraft und Beschleunigung, erhalten wir $ml\frac{\mathrm{d}^2\phi}{\mathrm{d}t^2} = -mg\sin(\phi(t))$, bzw. nach Umstellen

$$\frac{\mathrm{d}^2\phi}{\mathrm{d}t^2} + \frac{g}{l}\sin\phi(t) = 0.$$

Dies ist eine nichtlineare Differentialgleichung zweiter Ordnung, die die Bewegung bzw. Auslenkung $\phi(t)$ des Fadenpendels beschreibt. Die Gleichung besagt, dass die Winkelbeschleunigung des Fadenpendels in Richtung Ruhelage proportional zum Sinus des Auslenkungswinkels ist.

Differentialgleichungen sind Gleichungen, in denen eine Funktion und ihre Ableitungen vorkommen, und wir sehen hier ein Beispiel, wie interessante Funktionen in den Naturwissenschaften auftauchen: als Lösungen von Differentialgleichungen.

Im Fall des Fadenpendels ist die Differentialgleichung nicht durch eine elementare Funktion lösbar.

Alternativ kann die Bewegungsgleichung auch als Folgerung aus dem Energieerhaltungssatz gesehen werden. Die potentielle Energie des Fadenpendels ist $E_{\text{pot}} = mgh = mgl(1 - \cos\phi(t))$, wobei h die Höhe der Masse m über der Ruhelage ist (ebenfalls wieder mit einem geeigneten rechtwinkligen Dreieck zu sehen). Die kinetische Energie ist $E_{\text{kin}} = \frac{mv^2}{2} = \frac{m(l\phi'(t))^2}{2}$. Die Bewegungsgleichung ist dann

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}(E_{\text{pot}} + E_{\text{kin}}) = mgl\sin(\phi)\cdot\phi' + ml^2\phi'\phi'' = ml^2\phi'\left(\frac{g}{l}\sin\phi + \phi''\right)$$

Diese Gleichung ist erfüllt, wenn $\phi' = 0$ (keine Bewegung) oder wenn die oben hergeleitete Bewegungsgleichung erfüllt ist.

Näherung für kleine Auslenkungen: Taylorreihe für Sinus. Aus der Taylorreihe für den Sinus

$$\sin(\phi) = \phi - \frac{\phi^3}{6} + \frac{\phi^5}{120} - \cdots$$

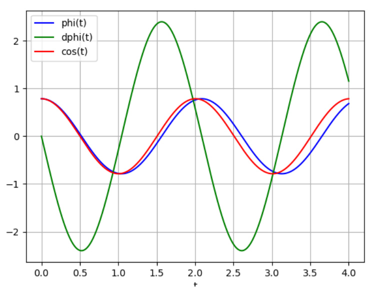
können wir für genügend kleine Winkel $\sin\phi \approx \phi$ annähern (s. Diskussion im Tutorium mit Fehlerabschätzung). Mit dieser Näherung vereinfacht sich die Gleichung zu $\phi''(t) + \frac{g}{l}\phi(t) = 0$. Dies ist dann eine lineare Differentialgleichung zweiter Ordnung und hat die allgemeine Lösung

$$\phi(t) = A\sin\left(\sqrt{\frac{g}{l}}t + \phi_0\right),$$

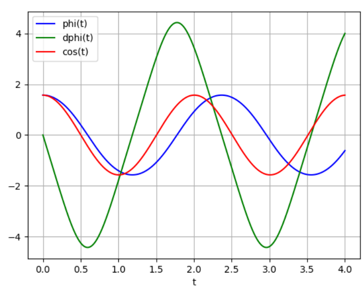
wobei A die maximale Auslenkung ist, und ϕ_0 der Auslenkungswinkel zum Zeitpunkt $t_0 = 0$. In dieser vereinfachten Situation kann die Periode des Fadenpendels bestimmt werden:

$$T_0 = 2\pi\sqrt{\frac{l}{g}}.$$

Die beiden folgenden Plots zeigen, dass die Näherung für große Auslenkungen nicht mehr besonders gut ist.



(A) Anfangswert $\phi = \pi/4$



(B) Anfangswert $\phi = \pi/2$

ABBILDUNG 30. Plots für $\phi(t)$, $\phi'(t)$ und $\cos(\sqrt{\frac{g}{l}}t)$

Potenzreihenansatz für Lösung der nichtlinearen DGL. Um die Bewegung des Fadenpendels zu beschreiben, müssen wir die Bewegungsgleichung $\phi'' + \frac{g}{l}\sin\phi = 0$ lösen. Verschiedene Methoden für solche Probleme werden in der Analysis 3 oder spezialisierteren Vorlesungen über Differentialgleichungen entwickelt.

Ein einfacher relativ allgemein verwendbarer Ansatz ist, eine Potenzreihendarstellung von $\phi(t)$ anzunehmen, und aus der Differentialgleichung Gleichungen abzuleiten, mit denen die Koeffizienten der Potenzreihe bestimmt werden können.

Wir schreiben als $\phi(t) = \sum_{k=0}^{\infty} a_k t^k$. Der Koeffizient a_0 beschreibt die Auslenkung zum Zeitpunkt $t_0 = 0$, der Koeffizient a_1 die Geschwindigkeit zum Zeitpunkt $t_0 = 0$. Wir wählen hier Anfangsauslenkung $a_0 = 0$ (t_0 ist also ein Durchgang durch die Ruhelage) und Anfangsgeschwindigkeit $a_1 = \frac{v_0}{l}$.

Nehmen wir an, dass die Potenzreihe für ϕ konvergiert, dann können wir die zweite Ableitung berechnen:

$$\phi''(t) = \sum_{k=2}^{\infty} k(k-1)a_k t^{k-2}.$$

Die Bewegungsgleichung ist dann

$$\sum_{k=2}^{\infty} k(k-1)a_k t^{k-2} + \frac{g}{l}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}\left(\sum_{k=0}^{\infty} a_k t^k\right)^{2n+1} = 0.$$

Hier benutzen wir, dass wir konvergente Potenzreihen (hier für ϕ) in konvergente Potenzreihen (hier für den Sinus) einsetzen können und das Ergebnis eine konvergente Potenzreihe ist, die wir durch Ausmultiplizieren bestimmen können. Schreiben wir die Potenzreihen aus, erhalten wir

$$\begin{aligned} & 2a_2 + 6a_3t + 12a_4t^2 + 20a_5t^3 + 30a_6t^4 + 42a_7t^5 + \cdots \\ + \quad & \frac{g}{l}\left(\frac{v_0}{l}t + a_2t^2 + a_3t^3 + a_4t^4 + \cdots\right) \\ - \quad & \frac{g}{3!l}\,(\text{Terme von Ordnung } \geq 3) + \cdots = 0 \end{aligned}$$

Durch Koeffizientenvergleich erhalten wir

$$a_2 = 0, \qquad a_3 = -\frac{gv_0}{6l^2}, \qquad a_4 = 0,$$

da in den Termen von $\sin\phi$ keine weiteren Terme von Ordnung ≤ 2 in t auftauchen. Diese Koeffizienten können wir wieder in die Bewegungsgleichung einsetzen:

$$\begin{aligned} & \left(-\frac{gv_0}{l^2}t + 20a_5t^3 + 30a_6t^4 + \cdots\right) + \frac{g}{l}\left(\frac{v_0}{l}t - \frac{gv_0}{6l^2}t^3 + a_5t^5 + \cdots\right) \\ - \quad & \frac{g}{3!l}\left(\frac{v_0^3}{l^3}t^3 - \frac{gv_0^3}{2l^4}t^5 + \cdots\right) + \frac{g}{5!l}\left(\frac{v_0^5}{l^5}t^5 + \cdots\right) = 0 \end{aligned}$$

Durch Koeffizientenvergleich können wir weitere Koeffizienten bestimmen:

$$a_5 = \frac{1}{120}\left(\frac{g^2v_0}{l^3} + \frac{gv_0^3}{l^4}\right), \qquad a_6 = 0, \qquad a_7 = -\frac{1}{5040}\left(\frac{g^3v_0}{l^4} + 11\frac{g^2v_0^3}{l^5} + \frac{gv_0^5}{l^6}\right)$$

Damit haben wir die ersten Terme der Potenzreihe für die Auslenkungsfunktion $\phi(t)$ bestimmt:

$$\phi(t) = \frac{v_0}{l}t - \frac{gv_0}{6l^2}t^3 + \frac{1}{120}\left(\frac{g^2v_0}{l^3} + \frac{gv_0^3}{l^4}\right)t^5 - \frac{1}{5040}\left(\frac{g^3v_0}{l^4} + 11\frac{g^2v_0^3}{l^5} + \frac{gv_0^5}{l^6}\right)t^7 + \cdots$$

Das Problem bei dieser Methode ist klar: Koeffizienten bestimmen ist schwierig, wir haben noch nicht einmal eine Restgliedabschätzung gemacht, um irgendwas über den Konvergenzradius zu sagen, mit der Potenzreihenform können auch schlecht Aussagen über die Periode der Funktion $\phi(t)$ gemacht werden....

Dies ist auch einer der Gründe, warum in der Analysis 2 Metriken und Topologien auf Funktionenräumen betrachtet werden: wir können versuchen, eine Lösung der Differentialgleichung als Grenzwert einer Funktionenfolge zu konstruieren. Dafür müssen wir natürlich aber über den Abstand zwischen Funktionen reden können. Damit können dann Aussagen darüber gemacht werden, wie gut eine Näherungslösung ist, oder wie schnell ein Näherungsverfahren gegen die tatsächliche Lösung konvergiert...

Periodendauer als elliptisches Integral. Wir können das Problem anders angehen, und den Energieerhaltungssatz als $E_{\text{pot}} + E_{\text{kin}} = E_0$ schreiben. Dann ist die Bewegungsgleichung

$$mgl(1-\cos(\phi)) + \frac{ml^2}{2}(\phi')^2 = E_0.$$

Umstellen liefert

$$\begin{aligned}
\frac{ml^2}{2}(\phi')^2 - mgl\cos(\phi) &= C := E_0 - mgl \\
(\phi')^2 &= \frac{2C}{ml^2} + \frac{2g}{l}\cos(\phi) \\
\frac{\phi'}{\sqrt{\frac{2C}{ml^2} + \frac{2g}{l}\cos(\phi)}} &= \pm 1.
\end{aligned}$$

Wir können nun die letzte Gleichung bezüglich t integrieren.

$$\int_{t_0}^{t_1} \frac{\phi'(t)\mathrm{d}t}{\sqrt{\frac{2C}{ml^2} + \frac{2g}{l}\cos(\phi)}} = \pm(t_1 - t_0)$$

Durch Substitution $\phi = \phi(t)$ (mit der Transformationsformel $\mathrm{d}\phi = \phi'\mathrm{d}t$ für die Differentiale) erhalten wir aus dem Integral bezüglich der Zeitvariablen t ein Integral bezüglich der Winkelvariablen ϕ.

$$\int_{\phi_0}^{\phi_1} \frac{\mathrm{d}\phi}{\sqrt{\frac{2C}{ml^2} + \frac{2g}{l}\cos(\phi)}} = \pm(t_1 - t_0).$$

In gewisser Weise beschreibt das Integral die Bewegung: geben wir uns die beiden Auslenkungen ϕ_0 und ϕ_1 vor, berechnet das Integral die Zeit, die vergeht, wenn sich das Fadenpendel von Auslenkung ϕ_0 zu Auslenkung ϕ_1 bewegt. Damit können wir insbesondere auch die Periodendauer formulieren: die halbe Periodendauer ist die Zeit, die benötigt wird, um von einer maximalen Auslenkung $-\phi_A$ zur anderen maximalen Auslenkung ϕ_A zu kommen:

$$T = 2\int_{-\phi_A}^{\phi_A} \frac{\mathrm{d}\phi}{\sqrt{\frac{2C}{ml^2} + \frac{2g}{l}\cos(\phi)}}.$$

Um den Integranden zu vereinfachen, benutzen wir

$$C = E_{\text{pot}} + E_{\text{kin}} - mgl = \frac{ml^2}{2}(\phi')^2 - mgl\cos\phi.$$

Im Punkt der maximalen Auslenkung ϕ_A ist die Geschwindigkeit $\phi' = 0$, also ist $C = -mgl\cos\phi_A$. Damit erhalten wir $\frac{2C}{ml^2} = -\frac{2mgl\cos\phi_A}{ml^2} = -\frac{2g}{l}\cos\phi_A$ und

$$T = 2\sqrt{\frac{l}{g}}\int_{-\phi_A}^{\phi_A} \frac{\mathrm{d}\phi}{\sqrt{2(\cos(\phi) - \cos(\phi_A))}}.$$

Mit Hilfe des Additionstheorems $\cos\phi = 1 - 2\sin^2\frac{\phi}{2}$ erhalten wir weiter

$$T = \sqrt{\frac{l}{g}}\int_{-\phi_A}^{\phi_A} \frac{\mathrm{d}\phi}{\sqrt{\sin^2\frac{\phi_A}{2} - \sin^2\frac{\phi}{2}}}.$$

Nun benutzen wir noch die Substitution

$$\sin\frac{\phi}{2} = \sin\frac{\phi_A}{2}\cdot\sin\frac{\theta}{2},$$

um das Periodenintegral in die Standardform der elliptischen Integrale zu bringen. Die Transformationsformel für die Differentiale ist wegen

$$\phi = 2\arcsin\left(\sin\frac{\phi_A}{2}\cdot\sin\frac{\theta}{2}\right)$$

dann

$$\mathrm{d}\phi = \frac{\sin\frac{\phi_A}{2}\cdot\cos\frac{\theta}{2}\mathrm{d}\theta}{\sqrt{1-\sin^2\frac{\phi_A}{2}\cdot\sin^2\frac{\theta}{2}}}.$$

Die Integrationsgrenzen $-\phi_A$ und ϕ_A gehen unter der Substitution auf $-\pi$ bzw. π, und wir erhalten

$$\begin{aligned} T &= \sqrt{\frac{l}{g}}\int_{-\phi_A}^{\phi_A} \frac{\mathrm{d}\phi}{\sqrt{\sin^2\frac{\phi_A}{2} - \sin^2\frac{\phi}{2}}} \\ &= \sqrt{\frac{l}{g}}\int_{-\pi}^{\pi} \frac{\sin\frac{\phi_A}{2}\cdot\cos\frac{\theta}{2}\mathrm{d}\theta}{\sqrt{1-\sin^2\frac{\phi_A}{2}\cdot\sin^2\frac{\theta}{2}}}\cdot\frac{1}{\sin\frac{\phi_A}{2}\sqrt{1-\sin^2\frac{\theta}{2}}} \\ &= \sqrt{\frac{l}{g}}\int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{\sqrt{1-\sin^2\frac{\phi_A}{2}\cdot\sin^2\frac{\theta}{2}}}. \end{aligned}$$

Das letzte Integral ist ein *elliptisches Integral der ersten Art*. (Elliptische Integrale verdanken ihren Namen der Tatsache, dass sie zuerst unter anderem im Zusammenhang mit Berechnungen des Umfangs von Ellipsen aufgetaucht sind.) Elliptische Integrale sind Standardbeispiele von Integralen, die nicht durch elementare Funktionen beschrieben werden können.

Auswertung elliptischer Integrale. Um jetzt Informationen über die Periodendauer des Fadenpendels zu bekommen, betrachten wir die elliptische Funktion

$$E(k) = \int_0^{\frac{\pi}{2}} \frac{\mathrm{d}t}{\sqrt{1-k^2\sin^2 t}}.$$

Die Periodendauer des Fadenpendels bei maximaler Auslenkung ϕ_A ist nach unseren vorigen Berechnungen

$$T = 4\sqrt{\frac{l}{g}}E\left(\sin\frac{\phi_A}{2}\right).$$

Der Graph der Funktion $E(k)$ für $0 \leq k < 1$ ist in Abbildung 31 zu sehen. Am Graphen können wir erkennen, dass die Differenz zwischen tatsächlicher Periode und der Näherung $2\pi\sqrt{\frac{l}{g}}$ für Auslenkungen nahe π beliebig groß werden können.

Um die Periodendauer jetzt näherungsweise zu berechnen, benutzen wir die Taylorreihe

$$\frac{1}{\sqrt{1-x^2}} = 1 + \frac{x^2}{2} + \frac{1\cdot 3}{2\cdot 4}x^4 + \cdots$$

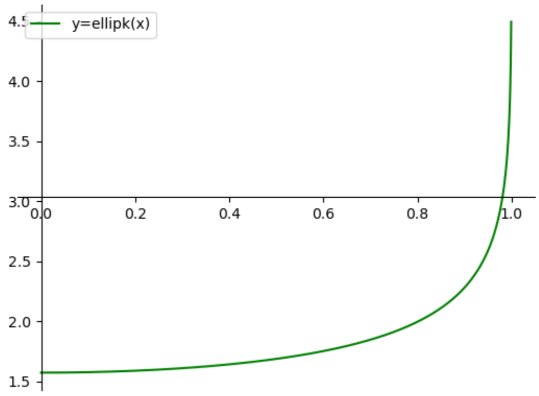


ABBILDUNG 31. Elliptische Funktion $E(k) = \int_0^{\frac{\pi}{2}} \frac{\mathrm{d}t}{\sqrt{1-k^2\sin^2 t}}$

Damit kann das elliptische Integral als Integral über eine konvergente Potenzreihe geschrieben werden, die dann gliedweise integriert und numerisch ausgewertet werden kann:

$$\begin{aligned} \int_0^\phi \frac{1}{\sqrt{1-k^2\sin^2 t}}\mathrm{d}t &= \int_0^\phi \left(1 + \frac{k^2\sin^2 t}{2} + \frac{1\cdot 3}{2\cdot 4}k^4\sin^4 t + \cdots\right)\mathrm{d}t \\ &= \phi + \frac{k^2}{2}\int_0^\phi \sin^2 t\mathrm{d}t + \frac{3k^4}{8}\int_0^\phi \sin^4 t\mathrm{d}t + \cdots \end{aligned}$$

Für die Integrale der Funktionen $\sin^{2n} t$ können die Rekursionsformeln aus Beispiel 7.5.3 benutzt werden.

KAPITEL 9

Metrische Räume und Topologie des $\mathbb{R}^n$

9.1. Normierte Vektorräume

Für die Analysis betrachten wir nur die Körper $K = \mathbb{R}$ oder $K = \mathbb{C}$.

Definition 9.1.1. *Sei $K = \mathbb{R}$ oder $\mathbb{C}$, und V ein K-Vektorraum. Eine* Norm *auf V ist eine Abbildung $\| - \|: V \to \mathbb{R}$, so dass für alle $v, w \in V$ und $\lambda \in K$ die folgenden Bedingungen erfüllt sind:*

positiv definit: $\|0\| = 0$ *und* $\|v\| > 0$ *für* $v \neq 0$.
absolut homogen: $\|\lambda \cdot v\| = |\lambda| \cdot \|v\|$
Dreiecksungleichung: $\|v + w\| \leq \|v\| + \|w\|$

Das Paar $(V, \| - \|)$ heißt auch normierter Vektorraum.

Beispiel 9.1.2. *(1) $\mathbb{R}$ und $\mathbb{C}$ sind normierte Vektorräume.*

*(2) K^n mit der p-*Norm*, $p \in \mathbb{R}, p \geq 1$:*

$$\|v\|_p := \sqrt[p]{\sum_{i=1}^{n} |v_i|^p}.$$

Der Spezialfall $n = 2$ ist die euklidische Norm*, der Fall $n = 1$ auch* Manhattan-Norm.

(3) Auf K^n haben wir die Maximumsnorm

$$\|(v_1, \ldots, v_n)\|_\infty = \max\{|v_1|, \ldots, |v_n|\}.$$

(4) Der Raum $C^0([a, b], \mathbb{R})$ der stetigen (reellwertigen) Funktionen auf dem Intervall $[a, b]$ mit der L^p-Norm, $p \in \mathbb{R}, p \geq 1$:

$$\|f\|_p := \sqrt[p]{\int_a^b |f(t)|^p \mathrm{d}t}$$

(5) Der Raum $C^0([a, b])$ der stetigen reellwertigen Funktionen auf dem Intervall $[a, b]$ mit der Supremumsnorm:

$$\|f\|_\infty := \sup\{|f(x)| \mid x \in [a, b]\}.$$

□

Eine Möglichkeit, eine Norm zu bekommen ist durch ein Skalarprodukt:

Definition 9.1.3. *Sei V ein K-Vektorraum mit $K = \mathbb{R}$ oder $K = \mathbb{C}$. Ein* Skalarprodukt *auf V ist eine positiv definite symmetrische Bilinearform, d.h. eine Abbildung $\langle -, - \rangle: V \times V \to \mathbb{R}$, so dass die folgenden Bedingungen erfüllt sind:*

Bilinearität: *Die Abbildung $\langle -, - \rangle$ ist $\mathbb{R}$-linear im ersten Argument, d.h. für alle $u, v, w \in V$ und alle $\lambda, \mu \in \mathbb{R}$ gilt*

$$\langle \lambda v + \mu w, u \rangle = \lambda \langle v, u \rangle + \mu \langle w, u \rangle$$

Symmetrie: *Für alle $v, w \in V$ gilt $\langle v, w \rangle = \overline{\langle w, v \rangle}$.*
positive Definitheit: *Für alle $v \in V$ gilt $\langle v, v \rangle \geq 0$, und $\langle v, v \rangle = 0$ genau dann, wenn $v = 0$.*

LEMMA 9.1.4 (Cauchy–Schwarz-Ungleichung). *Sei V ein $\mathbb{R}$-Vektorraum und $\langle -,-\rangle$ ein Skalarprodukt auf V. Für $v \in V$ definieren wir $\|v\| := \sqrt{\langle v,v\rangle}$. Dann gilt für alle Vektoren $v,w \in V$*

$$|\langle v,w\rangle| \leq \|v\| \cdot \|w\|.$$

BEWEIS. Im Fall $w = 0$ sind beide Seiten gleich 0, und die Ungleichung gilt.

Aus Bilinearität, Symmetrie und positiver Definitheit folgt für alle $v,w \in V$ und $\lambda \in \mathbb{R}$

$$\begin{aligned} 0 &\leq \langle v-\lambda w, v-\lambda w\rangle \\ &= \langle v-\lambda w, v\rangle - \lambda\langle v-\lambda w, w\rangle \\ &= \langle v,v\rangle - 2\lambda\langle v,w\rangle + \lambda^2\langle w,w\rangle. \end{aligned}$$

Wenn $w \neq 0$ ist, erhalten wir für den speziellen Wert $\lambda = \frac{\langle v,w\rangle}{\langle w,w\rangle}$ die Ungleichung

$$0 \leq \langle v,v\rangle - \frac{\langle v,w\rangle^2}{\langle w,w\rangle}$$

Umstellen und Wurzelziehen liefert die Behauptung. □

ÜBUNGSAUFGABE 9.1.5. *Sei V ein $\mathbb{R}$-Vektorraum mit einem Skalarprodukt $\langle -,-\rangle\colon V \times V \to \mathbb{R}$. Zeigen Sie, dass durch $\|v\| = \sqrt{\langle v,v\rangle}$ eine Norm auf V definiert wird.*

BEISPIEL 9.1.6. *Beispiele für Skalarprodukte (und entsprechend induzierte Normen):*

(1) Auf $\mathbb{R}^n$ wird durch

$$\langle (x_1,\ldots,x_n),(y_1,\ldots,y_n)\rangle = x_1y_1 + \cdots + x_ny_n$$

ein Skalarprodukt (Standardskalarprodukt) definiert. Die dazugehörige Norm ist die euklidische Norm $\|-\|_2$ aus Beispiel 9.1.2.

(2) Wir betrachten den Vektorraum der quadratsummierbaren reellen Folgen $(a_n)_n$, d.h. der Folgen $(a_n)_n$, für die $\sum_{n=0}^\infty a_n^2$ konvergiert. Auf diesem Vektorraum ist durch

$$\langle (a_n)_n,(b_n)_n\rangle = \sum_{n=0}^\infty a_nb_n$$

ein Skalarprodukt definiert.

(3) Auf dem komplexen Vektorraum der quadratintegrierbare Funktionen ist ein Skalarprodukt definiert:

$$\langle f,g\rangle = \int_a^b f(t)\overline{g(t)}\mathrm{d}t$$

Dieses Skalarprodukt spielt z.B. bei der Fourier-Transformation eine wichtige Rolle, s. Analysis 3.

□

9.2. Metrische Räume

DEFINITION 9.2.1. *Sei X eine Menge. Eine* Metrik *auf X ist eine Abbildung $d\colon X \times X \to \mathbb{R}$, so dass die folgenden Bedingungen erfüllt sind:*

Nicht-Negativität: *$d(x,x) = 0$ und $d(x,y) > 0$ für alle $x,y \in X$ mit $x \neq y$.*

Symmetrie: *$d(x,y) = d(y,x)$ für alle $x,y \in X$*

Dreiecksungleichung: *$d(x,y) \leq d(x,z) + d(z,y)$ für alle $x,y,z \in X$.*

Das Paar (X,d) heißt dann metrischer Raum.

BEISPIEL 9.2.2. *Beispiele für metrische Räume*

(1) *Sei* $(V, \|-\|)$ *ein normierter Vektorraum. Dann ist die folgende Abbildung eine Metrik auf* V*:*

$$d\colon V \times V \to \mathbb{R}\colon (x, y) \mapsto \|x - y\|$$

Spezialfälle sind die euklidische Metrik (induziert von $\|-\|_2$*) auf* $\mathbb{R}^n$ *oder* $\mathbb{C}^n$*.*

(2) *Auf der Menge der geordneten* n*-Tupel mit Einträgen* 0 *und* 1 *definiert der* Hamming-Abstand *eine Metrik. Der Hamming-Abstand zwischen zwei solchen Tupeln ist die dabei die Anzahl der Stellen, an denen zwei Tupel sich unterscheiden. Das ist relevant für Kodierungstheorie und die Konstruktion/Bewertung fehlerkorrigierender Codes.*

(3) *hyperbolische Metrik in der oberen Halbebene*

(4) *Graph: Anzahl der Kanten in kürzester Verbindung*

□

ÜBUNGSAUFGABE 9.2.3. *Zeigen Sie, dass die folgenden Abstandsbegriffe Metriken auf dem* $\mathbb{R}^n$ *definieren. (Dabei sind* $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ *und* $\|a\| = \sqrt{\sum a_i^2}$ *die euklidische Norm.)*

(1) *die Manhattan-Metrik* $d_{\mathrm{M}}(x, y) = \sum_{i=1}^n |x_i - y_i|$

(2) *die französische Eisenbahnmetrik für einen festgewählten Punkt* $p \in \mathbb{R}^n$*:*

$$d_{\mathrm{F},p}(x, y) = \begin{cases} \|x - y\| & x, y \text{ und } p \text{ liegen auf einer Geraden} \\ \|x - p\| + \|p - y\| & \text{sonst} \end{cases}$$

Berechnen Sie den Umfang des "Einheitskreises" $\{x \in \mathbb{R}^2 \mid d_{\mathrm{M}}(x, 0) = 1\}$ *in der Manhattanmetrik.*
Zeichnen Sie den "Einheitskreis" $\{x \in \mathbb{R}^2 \mid d_{\mathrm{F},P}(x, 0) = 1\}$ *in der französischen Eisenbahnmetrik, wobei der gewählte Punkt* $P = (0.5, 0)$ *ist.*

DEFINITION 9.2.4. *Sei* (X, d) *ein metrischer Raum. Der* offene Ball *von Radius* $r > 0$ *(offener* r*-Ball) um einen Punkt* $x \in X$ *ist definiert als*

$$\mathrm{B}(x; r) := \{y \in X \mid d(x, y) < r\}.$$

Der abgeschlossene Ball *von Radius* r *(abgeschlossener* r*-Ball) um einen Punkt* $x \in X$ *ist definiert als*

$$\overline{\mathrm{B}}(x; r) := \{y \in X \mid d(x, y) \leq r\}.$$

9.3. Topologische Räume

DEFINITION 9.3.1. *Ein* topologischer Raum *ist eine Menge* X *zusammen mit einer Menge* $\mathcal{O} \subseteq \mathcal{P}(X)$ *von Teilmengen von* X*, den sogenannten* offenen Mengen*, so dass die folgenden Bedingungen gelten:*

(1) $\emptyset$ *und* X *sind offen.*

(2) *Der Durchschnitt von endlich vielen offenen Mengen ist offen.*

(3) *Die Vereinigung von beliebig vielen offenen Mengen ist offen.*

Eine Menge $A \subseteq X$ *heißt* abgeschlossen*, wenn ihr Komplement offen ist.*

DEFINITION 9.3.2. *Sei* $(X, \mathcal{O})$ *ein topologischer Raum und* $x \in X$*. Eine offene Teilmenge* $U \in \mathcal{O}$ *mit* $x \in U$ *heißt* offene Umgebung von x*. Eine Teilmenge* $U \subseteq X$ *heißt* Umgebung von x*, wenn sie eine offene Umgebung von* x *enthält.*

Damit ist jede offene Menge eine Umgebung für jeden ihrer Punkte.

Beispiel 9.3.3. (1) *Sei X eine Menge. In der* diskreten *Topologie auf X sind alle Mengen offen. In der* indiskreten Topologie *auf X sind nur $\emptyset$ und X offen.*

Die kofinite Topologie auf X *hat als offene Mengen genau die Komplemente der endlichen Teilmengen in X und zusätzlich mit $\emptyset$ (wenn X unendlich ist).*

(2) *In der Standardtopologie/euklidischen Topologie auf $\mathbb{R}$ sind die offenen Mengen genau die Vereinigungen offener Intervalle.*

□

Allgemeiner wird für einen metrischen Raum (X, d) eine Topologie wie folgt definiert: eine Menge $U \subseteq X$ ist offen, wenn für jedes $x \in U$ ein $\epsilon > 0$ existiert, so dass die ϵ-Umgebung

$$\mathrm{B}(x; \epsilon) := \{y \in X \mid d(x, y) < \epsilon\}$$

in U enthalten ist. Dies definiert die *metrische Topologie*, ein Spezialfall ist insbesondere die "normale" euklidische Topologie auf $\mathbb{R}^n$.

Proposition 9.3.4. *Sei X ein metrischer Raum. Die metrische Topologie ist eine Topologie.*

Beweis. Es ist klar, dass $\emptyset$ und X offen sind.

Um zu zeigen, dass der Durchschnitt endlich vieler offener Mengen wieder offen ist, reicht es (Induktion), zu zeigen, dass der Durchschnitt von zwei offenen Mengen wieder offen ist. Seien U, V zwei offene Mengen und sei $x \in U \cap V$. Da U offen ist, existiert ein ϵ_U mit $\mathrm{B}(x; \epsilon_U) \subseteq U$; analog existiert ein ϵ_V mit $\mathrm{B}(x; \epsilon_V) \subseteq V$. Für $\epsilon = \min(\epsilon_U, \epsilon_V)$ gilt dann $\mathrm{B}(x; \epsilon) \subseteq U \cap V$, und damit ist $U \cap V$ offen.

Seien U_i, $i \in I$ offene Mengen, dabei ist die Indexmenge I beliebig. Für $x \in U := \bigcup_{i \in I} U_i$ existiert ein $i \in I$ mit $x \in U_i$ und (da U_i offen ist) ein ϵ mit $\mathrm{B}(x; \epsilon) \subseteq U_i \subseteq U$. Damit ist U offen. □

Übungsaufgabe 9.3.5. *Sei X ein metrischer Raum. Zeigen Sie, dass die offenen ϵ-Bälle $\mathrm{B}(x; \epsilon)$ offen in der metrischen Topologie sind. Zeigen Sie, dass die abgeschlossenen ϵ-Bälle $\overline{B}(x; \epsilon)$ abgeschlossen in der metrischen Topologie sind.*

Übungsaufgabe 9.3.6. (1) *Geben Sie zwei Beispiele an, dass der Schnitt von unendlich vielen offenen Mengen nicht unbedingt offen sein muss.*

(2) *Formulieren Sie eine Definition von topologischem Raum, die nur abgeschlossene Mengen benutzt und zeigen Sie, dass diese Definition äquivalent zu Definition 9.3.1 ist.*

Übungsaufgabe 9.3.7. *Wir betrachten $\mathbb{R}$ mit der Standardtopologie, d.h. offene Mengen sind genau die Vereinigungen offener Intervalle. Sind die folgenden Mengen offen bzw. abgeschlossen? Begründen Sie Ihre Aussage.*

$$\{0, 1, \sqrt{2}, \pi\}, \qquad \bigcup_{n=1}^{\infty} \left[0, 1 - \frac{1}{n}\right], \qquad \bigcup_{n=1}^{\infty} \left(n - \frac{1}{n}, n + \frac{1}{n}\right), \qquad \mathbb{Q}$$

Beispiel 9.3.8. *offene und abgeschlossene Mengen (und solche die es nicht sind) in $\mathbb{R}^n$* □

Definition 9.3.9. *Sei X ein topologischer Raum und $M \subseteq X$ eine Teilmenge.*

- *Das* Innere von M *ist die Menge*

$$\mathring{M} := \bigcup_{U \subseteq M, U \text{ offen}} U$$

- *Der* Abschluss von M *ist die Menge*

$$\overline{M} := \bigcap_{A \supseteq M, A \text{ abgeschlossen}} A$$

- *Der* Rand von M *ist die Menge* $\partial M := \overline{M} \setminus \mathring{M}$.
- *Wir sagen, dass* M dicht in X *ist, wenn* $\overline{M} = X$. *Wir sagen, dass* M nirgends dicht in X *ist, wenn* $\mathring{\overline{M}} = \emptyset$.
- *Ein Punkt* $p \in X$ *heißt* Häufungspunkt *von* M*, wenn in jeder Umgebung von* p *ein Punkt* $q \neq p$ *aus* M *liegt. Ein Punkt* $p \in X$ *heißt* Häufungspunkt von M im engeren Sinne*, wenn in jeder Umgebung von* p *unendlich viele Punkte aus* M *liegen.*

Übungsaufgabe 9.3.10. *Sei* X *ein metrischer Raum und* $M \subseteq X$ *eine Teilmenge. Zeigen Sie, dass ein Punkt* $p \in X$ *genau dann Häufungspunkt von* M *ist, wenn in jeder Umgebung von* p *unendlich viele Punkte aus* M *liegen.*

Übungsaufgabe 9.3.11. *Sei* X *ein topologischer Raum und* $A, B \subseteq X$ *Teilmengen. Zeigen Sie, dass aus* $A \subseteq B$ *auch* $\mathring{A} \subseteq \mathring{B}$ *und* $\overline{A} \subseteq \overline{B}$ *folgt. Geben Sie ein Beispiel an, dass zeigt, dass aus* $A \subseteq B$ *nicht unbedingt* $\partial A \subseteq \partial B$ *folgt.*

Beispiel 9.3.12. *Die rationalen Zahlen* $\mathbb{Q}$ *liegen dicht in* $\mathbb{R}$*. Ebenso liegt* $\mathbb{Q}^n$ *dicht in* $\mathbb{R}^n$ *(mit der Standardtopologie).*

Die Treppenfunktionen liegen dicht im Vektorraum der Regelfunktionen (mit der von der Supremumsnorm induzierten Topologie). □

Satz 9.3.13. *Sei* X *ein topologischer Raum, und* M *eine Teilmenge.*

(1) Bezeichnen wir mit $H(M)$ *die Menge der Häufungspunkte von* M*, dann ist* $\overline{M} = M \cup H(M)$*. Insbesondere ist* M *genau dann abgeschlossen, wenn alle Häufungspunkte von* M *schon in* M *enthalten sind.*

(2) Ein Punkt $p \in X$ *ist genau dann Randpunkt von* M*, d.h.* $p \in \partial M$*, wenn jede Umgebung von* p *sowohl Punkte aus* M *als auch aus* $X \setminus M$ *enthält.*

Beweis. (1) Wir zeigen $\overline{M} \subseteq M \cup H(M)$. Dafür zeigen wir, dass $M \cup H(M)$ abgeschlossen ist. Sei $M \subset X$ eine Menge und sei $x \in X \setminus M$ kein Häufungspunkt von M. Wir wollen zeigen, dass eine offene Umgebung U von x existiert, die keinen Punkt aus $M \cup H(M)$ enthält. Wenn in einer offenen Umgebung U von x ein Häufungspunkt h von M liegt, dann liegt auch ein Punkt $m \in M$ in U, da U eine offene Umgebung von h ist. Wenn also nun in jeder offenen Umgebung U von x ein Punkt aus $M \cup H(M)$ liegt, liegt auch in jeder offenen Umgebung ein Punkt $m \in M$, damit ist x selbst ein Häufungspunkt von M und wir erhalten einen Widerspruch. Also existiert eine offene Umgebung von x, die keinen Häufungspunkt von M enthält.

Nun zeigen wir $M \cup H(M) \subseteq \overline{M}$. Jede abgeschlossene Menge $A \subseteq X$, die M enthält, muss auch alle Häufungspunkte von M enthalten. Andernfalls existiert ein Häufungspunkt $x \in X \setminus A$. Da A abgeschlossen ist, ist $X \setminus A$ eine offene Umgebung von x, die keinen Punkt aus M enthält, was der Annahme widerspricht, dass x ein Häufungspunkt von M ist.

(2) Sei $p \in \partial M = \overline{M} \setminus \mathring{M}$. Wenn es eine offene Umgebung U von p mit $U \cap (X \setminus M) = \emptyset$ gibt, dann ist $U \subseteq M$ und damit $p \in \mathring{M}$, ein Widerspruch. Wenn es eine offene Umgebung U von p mit $U \cap M = \emptyset$ gibt, ist $p \notin M \cup H(M)$, ein Widerspruch.

Nehmen wir nun an, dass jede (offene) Umgebung von p Punkte von M enthält. Dann ist $p \in M \cup H(M) = \overline{M}$. Wenn nun jede offene Umgebung von p auch Punkte von $X \setminus M$ enthält, ist $p \notin \mathring{M}$. Daraus folgt $p \in \partial M$. □

Definition 9.3.14. *Sei X ein topologischer Raum. Dann heißt X* Hausdorff-Raum, *wenn für je zwei Punkte $x, y \in X$ offene Umgebungen U_x und U_y existieren, so dass $U_x \cap U_y = \emptyset$.*

Das bedeutet, dass Punkte durch disjunkte offene Mengen getrennt werden können. Dies ist zum Beispiel in der Konvergenztheorie für topologische Räume relevant für die Eindeutigkeit von Grenzwerten.

9.4. Konvergenz in topologischen Räumen

Definition 9.4.1. *Sei X ein topologischer Raum. Eine Folge $(x_n)_{n\in\mathbb{N}}$ von Punkten in X* konvergiert gegen den Grenzwert *$x \in X$, wenn für jede offene Umgebung U von x fast alle x_n in U liegen. Die Notation dafür ist wieder* $\lim_{n\to\infty} x_n = x$.

Beispiel 9.4.2.
(1) *Wir betrachten den Spezialfall $X = \mathbb{R}$ mit der vom Absolutbetrag induzierten Topologie. Eine offene Umgebung von $x \in \mathbb{R}$ ist eine Teilmenge $U \subseteq \mathbb{R}$, so dass ein ϵ existiert mit $(x - \epsilon, x + \epsilon) \subseteq U$. Umgekehrt ist auch jedes offene Intervall $(x - \epsilon, x + \epsilon)$ eine offene Umgebung von x. Nach Definition ist dann $\lim_{n\to\infty} x_n = x$, wenn für fast alle n gilt $|x_n - x| < \epsilon$. Das ist genau die Definition der Konvergenz aus der Analysis 1.*
(2) *Ebenso erhalten wir Konvergenz von komplexen Folgen als Spezialfall der obigen Definition, für die Topologie auf $\mathbb{C}$, die vom Absolutbetrag komplexer Zahlen induziert wird.*
(3) *Auch die gleichmäßige Konvergenz ist ein Spezialfall der obigen Konvergenzdefinition. In diesem Fall betrachten wir den topologischen Raum $C^0([a, b])$ mit der von der Supremumsnorm induzierten Topologie.*
(4) *Sei (X, d) ein metrischer Raum. Eine Folge $(x_n)_n$ in X konvergiert gegen $x \in X$ genau dann, wenn für alle $\epsilon > 0$ ein N existiert, so dass $d(x_n, x) < \epsilon$ für alle $n \geq N$ gilt.*

$\square$

Übungsaufgabe 9.4.3. *Sei X eine Menge mit der diskreten Topologie. Zeigen Sie, dass eine Folge $(x_n)_n$ genau dann gegen $x \in X$ konvergiert, wenn es ein $N \in \mathbb{N}$ gibt, so dass $x_n = x$ für alle $n \geq N$ gilt.*

Übungsaufgabe 9.4.4. *Sei X eine Menge mit der indiskreten Topologie. Zeigen Sie, dass jede Folge $(x_n)_n$ gegen jeden Punkt von X konvergiert.*

Zeigen Sie, dass in einem Hausdorff-Raum die Grenzwerte von konvergenten Folgen eindeutig sind.

Übungsaufgabe 9.4.5. *Sei X eine Menge. Welche Folgen konvergieren in der kofiniten Topologie auf X, und gegen welche Grenzwerte?*

Proposition 9.4.6. *Wir betrachten $\mathbb{R}^n$ mit der euklidischen Topologie. Eine Folge $(x_{1,k}, \ldots, x_{n,k})_{k\in\mathbb{N}}$ konvergiert genau dann, wenn die Folgen $(x_{i,k})_{k\in\mathbb{N}}$ für alle $1 \leq i \leq n$ konvergieren.*

Beweis. Wenn die Folge $(x_{1,k}, \ldots, x_{n,k})_{k\in\mathbb{N}}$ gegen den Grenzwert $(a_1, \ldots, a_n)$ konvergiert, dann ist die Folge der Abstände

$$\|x_k - a\|_2 = \sqrt{\sum_{i=1}^{n} (x_{i,k} - a_i)^2}$$

eine Nullfolge. Wegen $|x_{i,k} - a_i| \leq \sqrt{\sum_{i=1}^{n} (x_{i,k} - a_i)^2}$ folgt die komponentenweise Konvergenz.

Umgekehrt folgt aus der komponentenweisen Konvergenz der Folgen $(x_{i,k})_k$ gegen die Grenzwerte a_i, dass $|x_{i,k} - a_i|$ für alle k eine Nullfolge ist. Mit den Grenzwertrechenregeln folgt, dass die Folge $\sqrt{\sum_{i=1}^{n}(x_{i,k} - a_i)^2}$ eine Nullfolge ist, d.h. die Folge $(x_{1,k}, \ldots, x_{n,k})_{k\in\mathbb{N}}$ konvergiert gegen $(a_1, \ldots, a_n)$. □

BEMERKUNG 9.4.7. *Diese Argumentation funktioniert genauso für alle p-Normen und die Maximumsnorm.*

ÜBUNGSAUFGABE 9.4.8. *Welche der Folgen in $\mathbb{R}^n$ konvergieren?*

DEFINITION 9.4.9. *Sei X ein metrischer Raum. Eine Folge $(x_n)_n$ in X heißt* Cauchy-Folge, *wenn für jedes $\epsilon > 0$ ein N existiert, so dass für alle $n, m \geq N$ gilt $d(x_n, x_m) < \epsilon$.*

Ein metrischer Raum X heißt vollständig, *wenn jede Cauchy-Folge in X konvergiert.*

PROPOSITION 9.4.10. *Der $\mathbb{R}^n$ mit der euklidischen Metrik ist ein vollständiger metrischer Raum.*

BEWEIS. Sei $(x_{1,k}, \ldots, x_{n,k})_{k\in\mathbb{N}}$ eine Cauchy-Folge (für die euklidische Metrik). Dann gilt $|x_{i,k} - x_{i,l}| \leq \|(x_{1,k} - x_{1,l}, \ldots, x_{n,k} - x_{n,l})\|_2$, d.h. auch die Komponentenfolgen sind Cauchy-Folgen. Die Komponentenfolgen konvergieren wegen der Vollständigkeit von $\mathbb{R}$ aus der Analysis 1. Nach Proposition 9.4.6 konvergiert dann auch die Cauchy-Folge. □

BEISPIEL 9.4.11.
(1) *$\mathbb{Q}$ mit der Metrik $d(x, y) = |x-y|$ ist nicht vollständig. Cauchy-Folgen, deren Grenzwert irrational ist, konvergieren nicht in $\mathbb{Q}$.*
(2) *Der Vektorraum der Polynomfunktionen auf $[0, 1]$ mit der Supremumsnorm ist nicht vollständig. Für jede auf $[0, 1]$ konvergente Potenzreihe ist die Folge der Partialsummen eine Folge von Polynomen, deren Grenzwert aber im Allgemeinen kein Polynom ist. Zum Beispiel können wir die folgende Potenzreihe benutzen:*

$$\sum_{k=0}^{\infty} \frac{x^k}{2^k} = \frac{1}{1 - \frac{x}{2}}.$$

(3) *Der Raum $C^0([0, 1])$ mit der L^2-Norm ist ein weiteres Beispiel für einen normierten Vektorraum, der nicht vollständig ist. Es ist möglich, eine Folge $(f_n)_n$ zu finden, die eine Cauchy-Folge für die L^2-Norm ist und deren punktweiser Grenzwert die folgende unstetige Funktion ist:*

$$f(x) = \begin{cases} 0 & x \leq \frac{1}{2} \\ 1 & x > \frac{1}{2} \end{cases}$$

Ein ähnliches Beispiel lässt sich auch mit der L^1-Norm finden.

□

ÜBUNGSAUFGABE 9.4.12. *Sei I ein Intervall und $B(I)$ der Vektorraum der beschränkten stetigen Funktionen. Mit der Supremumsnorm wird $B(I)$ ein normierter Vektorraum. Zeigen Sie, dass der zugehörige metrische Raum vollständig ist.*

Hinweis: *Zeigen Sie erst gleichmäßige Konvergenz gegen den punktweisen Grenzwert. Zeigen Sie dann, dass der Grenzwert eine beschränkte stetige Funktion ist.*

DEFINITION 9.4.13. *Sei X ein metrischer Raum. Eine Teilmenge $M \subseteq X$ heißt* beschränkt, *wenn ein $x \in X$ und ein $R > 0$ existiert, so dass $M \subseteq \mathrm{B}(x; R)$ gilt. Eine Folge $(x_n)_n$ heißt beschränkt, wenn die Menge $\{x_n \mid n \in \mathbb{N}\}$ beschränkt ist.*

SATZ 9.4.14 (Bolzano–Weierstraß in $\mathbb{R}^n$). *Jede beschränkte Folge in $\mathbb{R}^n$ hat eine konvergente Teilfolge.*

Beweis. Die Aussage wird durch Induktion bewiesen, dabei ist der Induktionsanfang $n = 1$ der Satz von Bolzano–Weierstraß aus der Analysis 1. Für den Induktionsschritt sei $n \geq 2$ und wir setzen voraus, dass die Aussage für $\mathbb{R}^{n-1}$ gilt. Sei $(x_k)_k$ eine beschränkte Folge in $\mathbb{R}^n$. Dann ist das Bild der Folge $(x_k)_k$ unter der Projektionsabbildung $q\colon \mathbb{R}^n \to \mathbb{R}^{n-1}\colon (a_1, \ldots, a_n) \mapsto (a_1, \ldots, a_{n-1})$ ebenfalls wieder eine beschränkte Folge. Nach Induktionsvoraussetzung existiert also eine Teilfolge $(x_{j(k)})_{k\in\mathbb{N}}$, deren Bild unter q konvergent ist. Nach Bolzano–Weierstraß in $\mathbb{R}$ hat auch das Bild der Folge $(x_{j(k)})_{k\in\mathbb{N}}$ unter der Projektion $\mathrm{pr}_n\colon \mathbb{R}^n \to \mathbb{R}\colon (a_1, \ldots, a_n) \mapsto a_n$ eine konvergente Teilfolge. Damit existiert also eine Teilfolge $(x_{i(j(k))})_{k\in\mathbb{N}}$, deren Komponenten nach Konstruktion konvergieren. Damit haben wir eine konvergente Teilfolge, nach Proposition 9.4.6. □

Bemerkung 9.4.15. *Die Argumentation hier funktioniert auch für p-Normen und die Maximumnorm, da die relevante Proposition 9.4.6 für diese Normen gilt, s. die Bemerkung nach Proposition 9.4.6.*

Definition 9.4.16. *Sei V eine Vektorraum. Zwei Normen $\|-\|_1$ und $\|-\|_2$ auf V heißen* äquivalent, *wenn es Konstanten $\alpha, \beta > 0$ gibt, so dass für alle Vektoren $v \in V$ gilt*

$$\alpha\|v\|_1 \leq \|v\|_2 \leq \beta\|v\|_1.$$

Lemma 9.4.17. *Äquivalente Normen induzieren die gleiche Topologie.*

Beweis. Seien $\|-\|_1$ und $\|-\|_2$ zwei äquivalente Normen auf V, und seien $\mathcal{O}_1$ und $\mathcal{O}_2$ die entsprechenden induzierten metrischen Topologien. Wenn $U \in \mathcal{O}_1$, existiert nach Definition für jeden Vektor $v \in U$ ein ϵ, so dass $\mathrm{B}_1(v, \epsilon) = \{w \in V \mid \|v - w\|_1 < \epsilon\} \subseteq U$. Nach Voraussetzung ist $\alpha\|v - w\|_1 \leq \|v - w\|_2$. Das bedeutet

$$\mathrm{B}_2(v, \alpha\epsilon) = \{w \in V \mid \|v - w\|_2 < \alpha\epsilon\} \subseteq \{w \in V \mid \|v - w\|_1 < \epsilon\} = \mathrm{B}_1(v, \epsilon).$$

Damit ist also in jedem ϵ-Ball um v (für die Norm $\|-\|_1$) auch ein $\alpha\epsilon$-Ball um v für die Norm $\|-\|_2$ enthalten. Also ist $U \in \mathcal{O}_2$. Die umgekehrte Implikation wird analog mit der Ungleichung $\|v - w\|_2 \leq \beta\|v - w\|_1$ bewiesen. □

Satz 9.4.18. *Alle Normen auf einem endlich-dimensionalen Vektorraum sind äquivalent.*

Beweis. Da jeder endlich-dimensionale Vektorraum isomorph zu $\mathbb{R}^n$ ist, genügt es, die Aussage für $\mathbb{R}^n$ zu beweisen. Äquivalenz von Normen ist eine Äquivalenzrelation, also reicht es, zu zeigen, dass eine beliebige Norm äquivalent zur Maximumnorm ist. Sei $\|-\|$ eine Norm.

Wir bezeichnen mit e_i den i-ten Einheitsvektor, der als einzigen von 0 verschiedenen Eintrag eine 1 an der i-ten Stelle hat. Dann kann jeder Vektor $v \in \mathbb{R}^n$ als $v = \sum_{i=1}^n \lambda_i \mathrm{e}_i$ geschrieben werden, und es folgt

$$\|v\| \leq \sum_{i=1}^n |\lambda_i|\|\mathrm{e}_i\| \leq \beta\|v\|_\infty$$

mit $\beta = \|\mathrm{e}_1\| + \cdots + \|\mathrm{e}_n\|$.

Für die andere Abschätzung nehmen wir an, dass es kein $\alpha > 0$ gibt, so dass $\alpha\|v\|_\infty \leq \|v\|$ für alle $v \in \mathbb{R}^n$ gilt. Dann existiert zu jedem $k \in \mathbb{N}_{>0}$ ein v_k mit $\|v_k\|_\infty > k\|v_k\|$. Wir erhalten durch $w_k = \frac{v_k}{\|v_k\|_\infty}$ eine Folge von Vektoren $(w_k)_k$ in $\mathbb{R}^n$, für die gilt $\|w_k\| < \frac{1}{k}$. Nach Konstruktion ist $\|w_k\|_\infty = 1$, also ist die Folge beschränkt und besitzt damit nach Satz 9.4.14 eine konvergente Teilfolge. (Hier benutzen wir aber die Version des Satzes für die Maximumnorm, s. Bemerkung 9.4.15).

Wir haben also eine Folge $(u_k)_k$ konstruiert, die gegen einen Vektor u konvergiert, und damit gilt $\lim_{k\to\infty} \|u_k - u\|_\infty = 0$. Wir können jetzt schreiben

$$\|u\| = \|u - u_k + u_k\| \le \|u - u_k\| + \|u_k\| \le \beta\|u - u_k\|_\infty + \frac{1}{k}$$

Die rechte Seite konvergiert gegen 0, also ist $\|u\| = 0$, damit $u = 0$. Andererseits ist $1 = \|u_k\|_\infty = \|u - (u_k - u)\|_\infty \le \|u\|_\infty + \|u_k - u\|_\infty$, und wir erhalten einen Widerspruch. □

BEMERKUNG 9.4.19. *Das bedeutet, dass alle Normen (insbesondere p-Normen für $p \in \mathbb{R}$, $p \ge 1$ und Maximumnorm) die gleiche metrische Topologie auf $\mathbb{R}^n$ induzieren. Für die Analysis bedeutet das, dass alle topologischen Aussagen über $\mathbb{R}^n$ mit beliebigen Normen bewiesen werden können - Abschätzungen können also immer mit der für die Fragestellung am Besten geeigneten Norm durchgeführt werden. Das ist nicht nur zum Rechnen und Beweisen praktisch, sondern auch konzeptionell zentral: Normen, die nicht die gleiche Topologie induzieren, hätten dann auch unterschiedliche Konvergenzbegriffe, und könnten - abhängig von der Norm bzw. Topologie - gegen unterschiedliche Grenzwerte konvergieren. Bei unendlich-dimensionalen Vektorräumen kann das auch passieren: so ist zum Beispiel $C^0([0,1])$ bezüglich der Supremumsnorm vollständig, aber bezüglich der L^1-Norm nicht vollständig.*

ÜBUNGSAUFGABE 9.4.20. *Geben Sie ein Beispiel einer Folge $(f_n)_n$ in $C^0([0,1])$ an, die für die L^1-Norm eine Cauchy-Folge ist, aber für die Supremumsnorm keine Cauchy-Folge ist.*

9.5. Stetige Abbildungen

DEFINITION 9.5.1 (stetige Abbildung). *Seien X, Y zwei topologische Räume. Eine Abbildung $f\colon X \to Y$ (von Mengen) heißt* stetig im Punkt $p \in X$, *wenn es zu jeder Umgebung V von $f(p)$ eine Umgebung U von p gibt, so dass $f(U) \subseteq V$ (oder äquivalent $U \subseteq f^{-1}(V)$) gilt. Eine Abbildung $f\colon X \to Y$ heißt* stetige Abbildung, *wenn f an allen Punkten $p \in X$ stetig ist.*

ÜBUNGSAUFGABE 9.5.2. *Zeigen Sie, dass eine Abbildung $f\colon X \to Y$ genau dann stetig ist, wenn für alle offenen Mengen $U \subseteq Y$ das Urbild $f^{-1}(U)$ offen in X ist.*

BEISPIEL 9.5.3. *Für alle topologischen Räume X, Y sind konstante Abbildungen $f\colon X \to Y$ stetig.* □

SATZ 9.5.4 (ϵ-δ-Kriterium für Stetigkeit). *Seien (X, d_X) und (Y, d_Y) metrische Räume. Dann ist eine Abbildung $f\colon X \to Y$ genau dann stetig in $p \in X$, wenn für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so dass $d_Y(f(x), f(p)) < \epsilon$ für alle $x \in X$ mit $d_X(x,p) < \delta$ gilt.*

BEWEIS. Sei $f\colon X \to Y$ stetig im Punkt $p \in X$, und sei $\epsilon > 0$. Dann ist $\mathrm{B}(f(p);\epsilon) = \{y \in Y \mid d_Y(y, f(p)) < \epsilon\}$ eine offene Umgebung von $f(p)$, und da f stetig ist, existiert eine Umgebung U von p mit $U \subseteq f^{-1}(\mathrm{B}(f(p);\epsilon))$. Nach Definition der metrischen Topologie auf X existiert dann ein δ mit

$$\{x \in X \mid d_X(x,p) < \delta\} = \mathrm{B}(p,\delta) \subseteq U \subseteq f^{-1}(\mathrm{B}(f(p);\epsilon)).$$

Das ist aber genau das ϵ-δ-Kriterium.

Sei nun $f\colon X \to Y$ so, dass das ϵ-δ-Kriterium erfüllt ist, und sei $V \subseteq Y$ eine Umgebung von $f(p)$. Nach Definition existiert dann ein $\epsilon > 0$, so dass $\mathrm{B}(f(p);\epsilon) \subseteq V$. Nach dem ϵ-δ-Kriterium existiert ein $\delta > 0$, so dass $\mathrm{B}(p;\delta) \subseteq f^{-1}(\mathrm{B}(f(p);\epsilon)) \subseteq f^{-1}(V)$. Mit $U = \mathrm{B}(p;\delta)$ sehen wir die Stetigkeit von f am Punkt p. □

Beispiel 9.5.5. *Eine Funktion $f\colon \mathbb{R} \to \mathbb{R}$, die stetig im Sinne der Definition 4.1.1 aus Analysis 1 ist, ist stetig im Sinne der Definition 9.5.1 für die Standardtopologie auf $\mathbb{R}$.* □

Definition 9.5.6. *Eine Abbildung $f\colon X \to Y$ zwischen metrischen Räumen heißt* Lipschitz-stetig, *wenn es eine Konstante $L \geq 0$ gibt, so dass $d_Y(f(x), f(y)) < Ld_X(x, y)$ für alle $x, y \in X$ gilt.*

Beispiel 9.5.7. *(1) Lipschitz-stetige Abbildungen sind stetig.*

(2) Lineare Abbildungen $f\colon \mathbb{R}^n \to \mathbb{R}^m$. Wir benutzen dafür die Maximumnorm $\|-\|_\infty$ auf $\mathbb{R}^n$ und eine beliebige Norm $\|-\|$ auf $\mathbb{R}^m$. Schreiben wir $x, y \in \mathbb{R}^n$ in der Standardbasis als $x = \sum_{i=1}^n x_i \mathrm{e}_i$ bzw. $y = \sum_{i=1}^n y_i \mathrm{e}_i$, ist

$$\|f(x) - f(y)\| = \left\| \sum_{i=1}^n (x_i f(\mathrm{e}_i) - y_i f(\mathrm{e}_i)) \right\| \leq \sum_{i=1}^n |x_i - y_i| \|f(\mathrm{e}_i)\| \leq L\|x - y\|_\infty.$$

Die lineare Abbildung f ist dann Lipschitz-stetig mit Lipschitz-Konstante $L = \sum_{i=1}^n \|f(\mathrm{e}_i)\|$.

(3) Jede Norm $\|-\|\colon V \to \mathbb{R}$ ist Lipschitz-stetig mit Lipschitz-Konstante 1, da $|\|x\| - \|y\|| \leq \|x - y\|$.

□

Proposition 9.5.8. *(1) Gegeben seien Abbildungen $f\colon X \to Y$ und $g\colon Y \to Z$. Wenn f stetig in $p \in X$ ist und g stetig in $f(p)$ ist, dann ist auch $g \circ f\colon X \to Z$ im Punkt p stetig.*

(2) Sei $K = \mathbb{R}$ oder $K = \mathbb{C}$ (jeweils mit der vom Absolutbetrag induzierten Topologie). Für Abbildungen $f, g\colon X \to K$, die im Punkt $p \in X$ stetig sind, sind auch $f + g$ und $f \cdot g$ in p stetig. Wenn $g(p) \neq 0$, ist auch $\frac{f}{g}$ stetig in p.

Beispiel 9.5.9. *Damit sind Polynome und rationale Funktionen in mehreren Variablen stetig auf ihrem Definitionsbereich.* □

Bemerkung 9.5.10. *Sei $f\colon X \to \mathbb{R}$ eine stetige Abbildung. Für jede reelle Zahl $c \in \mathbb{R}$ ist die Subniveaumenge $\{x \in X \mid f(x) < c\}$ offen, und die Subniveaumenge $\{x \in X \mid f(x) \leq c\}$ abgeschlossen. Die Niveaumenge $f^{-1}(c) = \{x \in X \mid f(x) = c\}$ ist abgeschlossen.*

Bemerkung 9.5.11. *Für eine stetige Abbildung $f\colon X \to Y$ ist das Bild $f(U)$ einer offenen (bzw. abgeschlossenen) Menge $U \subseteq X$ nicht notwendigerweise offen (bzw. abgeschlossen) in Y.*

(1) Betrachten wir zum Beispiel die Abbildung $\mathbb{R}^2 \to \mathbb{R}\colon (x, y) \to x$. Der Graph der Funktion $f(x) = \frac{1}{x}$ ist als Niveaumenge $\{(x, y) \in \mathbb{R}^2 \mid xy = 1\}$ der stetigen Funktion $f(x, y) = xy$ abgeschlossen in $\mathbb{R}^2$. Das Bild des Graphen in $\mathbb{R}$ ist $\mathbb{R} \setminus \{0\}$, also nicht abgeschlossen.

(2) Die konstante Abbildung $f\colon X \to \mathbb{R}\colon x \mapsto 0$ bildet jede nichtleere offene Menge in X auf die Menge $\{0\}$ ab, die nicht offen ist.

Definition 9.5.12. *Seien X, Y topologische Räume, Y ein Hausdorff-Raum und $D \subseteq X$ eine Teilmenge. Eine Abbildung $f\colon D \to Y$ hat in einem Häufungspunkt $a \in H(D) \subseteq X$ von D den Grenzwert $b \in Y$, wenn für jede Umgebung V von $b \in Y$ eine Umgebung U von $a \in X$ existiert, so dass $f(U \cap D) \subseteq V$. Die Notation ist $\lim_{x \to a} f(x) = b$.*

Proposition 9.5.13. *Seien X, Y topologische Räume, Y ein Hausdorff-Raum und $D \subseteq X$ eine Teilmenge. Eine Abbildung $f\colon D \to Y$ ist genau dann stetig in einem Häufungspunkt $p \in D$, wenn $\lim_{x \to p} f(x) = f(p)$ gilt.*

Beispiel 9.5.14. *Sei $[a,b] \subseteq \mathbb{R}$ ein abgeschlossenes Intervall. Auf dem Vektorraum der Regelfunktionen $\mathcal{R}([a,b])$ mit der von der Supremumsnorm induzierten Topologie ist das Integral*

$$\int : \mathcal{R}([a,b]) \to \mathbb{R} \colon f \mapsto \int_a^b f(t)\mathrm{d}t$$

eine stetige Abbildung. Dies folgt aus Proposition 9.5.13 und Proposition 7.7.3: die Folgenstetigkeit aus Proposition 9.5.13 ist genau die Aussage aus Proposition 7.7.3, dass das Integral mit gleichmäßig konvergenten Grenzwerten vertauscht. Da die Treppenfunktionen $\mathcal{T}([a,b])$ eine dichte Teilmenge in $\mathcal{R}([a,b])$ ist das Integral als stetige Funktion schon durch die Werte auf den Treppenfunktionen eindeutig bestimmt. Das bedeutet, es gibt nur eine Möglichkeit, eine stetige reellwertige Abbildung auf Regelfunktionen zu definieren, die auf Treppenfunktionen den Flächeninhalt unter dem Graphen liefert – das Regelintegral. □

Definition 9.5.15. *Eine bijektive stetige Abbildung $f \colon X \to Y$ von topologischen Räumen heißt* Homöomorphismus, *wenn die Umkehrabbildung $f^{-1} \colon Y \to X$ stetig ist.*

Beispiel 9.5.16. *Die Abbildung*

$$f \colon \mathrm{B}(0,1) \subseteq \mathbb{R}^n \to \mathbb{R}^n \colon x \mapsto \frac{x}{1 - \|x\|}$$

ist ein oft benutzter Homöomorphismus mit Umkehrabbildung $f^{-1}(x) = \frac{x}{1+\|x\|}$. Aus Sicht der Topologie gibt es keinen Unterschied zwischen $\mathbb{R}^n$ und $\mathrm{B}(0,1)$. Ebenso folgt, dass Einheitsbälle für beliebige Normen auf $\mathbb{R}^n$ homöomorph sind. □

Definition 9.5.17 (Unterraum). *Sei $(X, \mathcal{O}(X))$ ein topologischer Raum und $Y \subseteq X$ eine Teilmenge von X. Dann definiert $\mathcal{O}(Y) = \{Y \cap U \mid U \in \mathcal{O}(X)\}$ eine Topologie auf Y, und $(Y, \mathcal{O}(Y))$ heißt* Unterraum von X. *Wenn $Y \in \mathcal{O}(X)$ ist, spricht man von* offenem Unterraum, *wenn $X \setminus Y \in \mathcal{O}(X)$ von* abgeschlossenem Unterraum.

Beispiel 9.5.18. *Die von $\mathbb{C}$ induzierte Topologie auf $\mathbb{R}$ ist die Standardtopologie.* □

Definition 9.5.19. *Für topologische Räume X, Y ist die* Produkttopologie *auf $X \times Y$ erzeugt durch die Menge*

$$\{U \times V \mid U \text{ offen in } X \text{ und } V \text{ offen in } Y\}$$

der "offenen Rechtecken", d.h. eine Menge ist genau dann offen, wenn sie Vereinigung von offenen Rechtecken ist. Alternativ ist eine Menge $U \subseteq X \times Y$ genau dann offen, wenn jeder Punkt $x \in U$ in einem offenen Rechteck enthalten ist, das selbst ganz in U enthalten ist.

Beispiel 9.5.20. *Auf $\mathbb{R}^n \cong \prod_{i=1}^n \mathbb{R}$ haben wir die Produkttopologie, wobei wir auf den Faktoren $\mathbb{R}$ die Standardtopologie benutzen. Diese Produkttopologie stimmt mit der Standardtopologie auf $\mathbb{R}^n$ überein. Das sieht man am Besten mit der Maximumnorm, deren offene Bälle offene Rechtecke (sogar Würfel) in der Produkttopologie sind.* □

9.6. Kompaktheit

Definition 9.6.1. *Sei X ein topologischer Raum, $K \subseteq X$ eine Teilmenge. Eine* offene Überdeckung *von K ist eine Familie offener Mengen $\{U_i\}_{i \in I}$ mit $K \subseteq \bigcup_{i \in I} U_i$.*

Die Teilmenge K heißt kompakt, *wenn jede offene Überdeckung eine endliche Teilüberdeckung hat, d.h. für jede offene Überdeckung $\{U_i\}_{i\in I}$ existiert eine endliche Teilmenge $J \subseteq I$ mit $X = \bigcup_{j\in J} U_j$.*

Beispiel 9.6.2. *Beispiele für Überdeckungen:*

- *Es gibt viele verschiedene Möglichkeiten, eine offene Überdeckung von $\mathbb{R}$ durch offene Intervalle zu erhalten: $\{(-i,i) \mid i \in \mathbb{Z}\}$ ist eine Möglichkeit; $\{(i-1,i+1) \mid i \in \mathbb{Z}\}$ eine andere; es gibt noch viele weitere...*
- *Analog können wir $\mathbb{R}^n$ durch offene ϵ-Bälle überdecken. Eine einfache Möglichkeit ist $\{\mathrm{B}(0,i) \mid i \in \mathbb{Z}\}$ (funktioniert für jede Norm). Eine andere Möglichkeit sind r-Bälle um die Gitterpunkte $\mathbb{Z}^n$, also die Überdeckung $\{\mathrm{B}(i,r) \mid i \in \mathbb{Z}^n\}$. Dabei muss der Radius r so groß gewählt werden, dass tatsächlich jeder Punkt $p \in \mathbb{R}^n$ in einem solchen Ball liegt. Der Radius muss also größer als der Abstand zum nächstliegenden Gitterpunkt sein - das wäre in $\mathbb{R}^2$ zum Beispiel 1 für die Manhattan-Norm, $\frac{\sqrt{2}}{2}$ für die euklidische Norm und $\frac{1}{2}$ für die Maximumsnorm.*

□

Beispiel 9.6.3. *Beispiele für kompakte bzw. nicht-kompakte Mengen:*

- *Die indiskrete Topologie ist immer kompakt, da es nur eine nichtleere offene Menge gibt. Die diskrete Topologie auf einer Menge M ist genau dann kompakt, wenn M endlich viele Elemente enthält, da $\{\{m\} \mid m \in M\}$ eine offene Überdeckung ist.*
- *In $\mathbb{R}$ mit der Standardtopologie sind abgeschlossene Intervalle kompakt, s. Satz von Heine–Borel. Für die Überdeckung von $\mathbb{R}$ durch $V_i = (i-1,i+1)$, $i \in \mathbb{Z}$ ist $(-3,-1),(-2,0),(-1,1)$ eine endliche Teilüberdeckung für das Intervall $[-2,0]$.*

  Offene Intervalle sind dagegen nicht kompakt: Für das Intervall $(-1,1)$ ist zum Beispiel $\{(-\epsilon,\epsilon) \mid 0 < \epsilon < 1\}$ eine offene Überdeckung, die keine endliche Teilüberdeckung hat.

□

Definition 9.6.4. *Sei X ein topologischer Raum X. Eine Teilmenge $M \subseteq X$ heißt folgenkompakt, wenn jede Folge in M eine konvergente Teilfolge besitzt, deren Grenzwert in M liegt.*

Satz 9.6.5. *Sei X ein metrischer Raum. Dann gilt*

(1) Wenn X kompakt ist, dann ist X auch folgenkompakt.
(2) Jede folgenkompakte Teilmenge $M \subseteq X$ ist abgeschlossen und beschränkt.

Beweis. (1) Für eine Folge $(x_k)_k$, betrachten wir die Menge $M = \{x_k \mid k \in \mathbb{N}\}$. Wir können annehmen, dass M unendlich ist (andernfalls hat $(x_k)_k$ eine konstante Teilfolge). Nehmen wir an, dass M keinen Häufungspunkt hat. Dann existiert für jeden Punkt $x \in X$ eine offene Umgebung U_x, so dass $U_x \cap M$ höchstens x selbst enthält. Dann ist $\{U_x \mid x \in X\}$ eine offene Überdeckung von X. Da X kompakt ist, gibt es eine endliche Teilüberdeckung $U_{x_1},\ldots,U_{x_r}$. Dies widerspricht der Annahme, dass M unendlich ist.

Sei nun $p \in X$ ein Häufungspunkt von M. Wir konstruieren eine Teilfolge, die gegen p konvergiert. Für jedes $n \in \mathbb{N}$ enthält der Ball $\mathrm{B}(p,\frac{1}{n})$ einen Punkt $x_{k(n)} \in M$ mit $x_{k(n)} \neq p$. Nach Konstruktion konvergiert diese Teilfolge von $(x_k)_k$ gegen p.

(2) Sei $K \subseteq X$ folgenkompakt. Nehmen wir an, dass K nicht beschränkt ist. Für einen gewählten Punkt $p \in X$ existiert dann für jedes n ein Punkt $x_n \in K$ mit

$d(x_n, p) > n$. Nach Konstruktion kann die Folge keine konvergente Teilfolge haben, was der Folgenkompaktheit widerspricht.

Sei p ein Häufungspunkt von K. Wir können nun wie in (1) eine Folge $(x_n)_n$ von Punkten in K konstruieren, die gegen p konvergiert. Nach Voraussetzung existiert eine konvergente Teilfolge, die gegen einen Punkt $q \in K$ konvergiert. Aber Grenzwerte in metrischen Räumen sind eindeutig, also ist $p \in K$, und damit ist K abgeschlossen. □

Satz 9.6.6 (Satz von Heine–Borel). *Sei $K \subseteq \mathbb{R}^n$ eine Teilmenge. Dann sind die folgenden Aussagen äquivalent:*

(1) K ist abgeschlossen und beschränkt.
(2) K ist kompakt.
(3) K ist folgenkompakt.

Beweis. Es bleibt nach Satz 9.6.5 nur noch die Implikation (1)⇒(2) zu zeigen. Sei $U_i, i \in I$ eine offene Überdeckung von K, die keine endliche Teilüberdeckung hat. Da K beschränkt ist, liegt K in einem Würfel W_0 der Kantenlänge $s > 0$. Wir können W_0 in 2^n Teilwürfel mit Kantenlänge $\frac{s}{2}$ zerlegen. Da W_0 nicht von endlich vielen U_i überdeckt wird, gibt es auch einen der kleineren Würfel W_1, so dass $K \cap W_1$ nicht von endlich vielen U_i überdeckt wird. Iterativ erhalten wir eine Folge von Würfeln $W_0 \supset W_1 \supset W_2 \supset \cdots$, der Kantenlänge $\frac{s}{2^k}$, so dass keine der Mengen $K \cap W_k$ von endlich vielen U_i überdeckt wird. Für jedes k ist damit $K \cap W_k$ nicht leer, und wir können einen Punkt $x_k \in K \cap W_k$ wählen. Nach Konstruktion ist $(x_k)_k$ eine Cauchy-Folge, die gegen einen Grenzwert $p \in K$ konvergiert, da K abgeschlossen ist. Da K von den U_i überdeckt wird, existiert also eine offene Menge U_i mit $p \in U_i$. Dann sind aber schon fast alle Würfel W_k in U_i enthalten, was der Annahme widerspricht. Also ist K kompakt. □

Bemerkung 9.6.7. *In Funktionenräumen gibt es auch Beispiele für abgeschlossene beschränkte Mengen, die nicht kompakt sind. Zum Beispiel die Einheitskugel* $\mathrm{B}(0,1)$ *in* $C^0([0,1],\mathbb{C})$ *mit der* L^2*-Norm.*

Jetzt können wir Konsequenzen für stetige Abbildungen formulieren:

Satz 9.6.8. *Sei $f\colon X \to Y$ eine stetige Abbildung. Für eine kompakte Teilmenge $K \subseteq X$ ist auch das Bild $f(K)$ kompakt.*

Beweis. Sei U_i, $i \in I$ eine offene Überdeckung von $f(K)$. Dann ist $f^{-1}(U_i)$ eine offene Überdeckung von K, da f stetig ist. Da K kompakt ist, existiert eine endliche Teilüberdeckung $f^{-1}(U_{i_1}), \ldots, f^{-1}(U_{i_n})$ von K, aber damit ist $U_{i_1}, \ldots, U_{i_n}$ eine endliche Teilüberdeckung von $f(K)$. □

Korollar 9.6.9. *Sei $X \neq \emptyset$ ein nicht-leerer kompakter topologischer Raum und $f\colon X \to \mathbb{R}$ eine stetige Abbildung. Dann hat f in X ein Maximum und ein Minimum.*

Beweis. Nach Satz 9.6.8 ist $f(X)$ kompakt, nach dem Satz von Heine–Borel 9.6.6 also beschränkt und abgeschlossen. Aus der Beschränktheit folgt, dass $f(X)$ ein Infimum und Supremum hat. Infimum und Supremum Häufungswerte sind von $f(X)$. Da $f(X)$ auch abgeschlossen ist, sind Infimum und Supremum in $f(X)$ enthalten und sind dann Minimum bzw Maximum der Funktionswerte. □

Proposition 9.6.10. *Sei $f\colon X \to Y$ eine stetige Abbildung zwischen metrischen Räumen, und sei X kompakt. Dann ist f gleichmäßig stetig, d.h. für jedes $\epsilon > 0$ existiert ein $\delta > 0$, so dass für alle $x, y \in X$ mit $d_X(x,y) < \delta$ gilt $d_Y(f(x), f(y)) < \epsilon$.*

9.7. Zusammenhang

DEFINITION 9.7.1. *Ein topologischer Raum X heißt* zusammenhängend*, wenn es keine Zerlegung $X = U \cup V$ in zwei nicht-leere, offene und disjunkte Teilmengen U, V gibt. Eine Teilmenge $M \subseteq X$ heißt* zusammenhängend*, wenn sie in der induzierten Teilraum-Topologie zusammenhängend ist.*

SATZ 9.7.2. *Eine Teilmenge $M \subseteq \mathbb{R}$ ist genau dann zusammenhängend, wenn sie ein Intervall ist.*

BEWEIS. □

SATZ 9.7.3. *Sei $f: X \to Y$ eine stetige Abbildung. Für eine zusammenhängende Teilmenge $M \subseteq X$ ist das Bild $f(M)$ wieder zusammenhängend.*

BEWEIS. □

KOROLLAR 9.7.4 (Zwischenwertsatz). *Sei X ein zusammenhängender topologischer Raum und $f: X \to \mathbb{R}$ stetig. Für zwei Punkte $a, b \in X$ nimmt f jeden Wert zwischen $f(a)$ und $f(b)$ an.*

Bemerkung zu Wegzusammenhang

KAPITEL 10

Differentialrechnung in mehreren Variablen

10.1. Vektorwertige Abbildungen in mehreren Veränderlichen

Für die Differentialrechnung wollen wir nun allgemein Abbildungen $f\colon D \to \mathbb{R}^m$ mit Definitionsbereich $D \subseteq \mathbb{R}^n$ betrachten und für solche Funktionen die Methoden und Aussagen der Differentialrechnung weiterentwickeln. Allgemein sprechen wir von Funktionen *in mehreren Veränderlichen* oder *mehreren Variablen*, wenn $D \subseteq \mathbb{R}^n$ (da die Funktion in diesem Fall von n Argumenten abhängt). Funktionen der Form $f\colon D \to \mathbb{R}^m$ werden auch *vektorwertige Funktionen* genannt.

BEISPIEL 10.1.1. *Eine Funktion $\gamma\colon D \to \mathbb{R}^n$ mit $D \subseteq \mathbb{R}$ heißt auch* Kurve. *Wenn $n = 2$ ist, sprechen wir von* ebenen Kurven, *für $n = 3$ von* Raumkurven. *Ein Standard-Anwendungsfall für Kurven ist die Modellierung von Bewegungen.*

Die Kurve

$$\gamma\colon \mathbb{R} \to \mathbb{R}^2\colon t \mapsto (\cos t, \sin t)$$

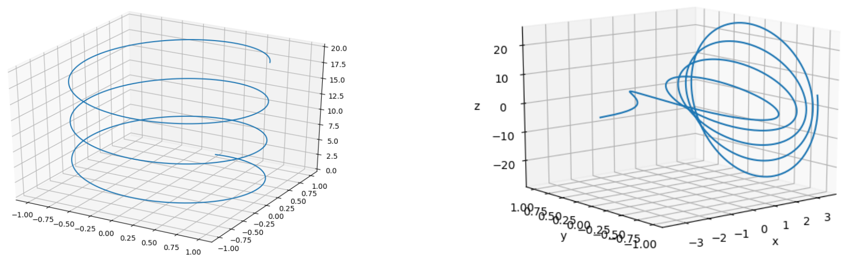
beschreibt eine Bewegung auf dem Einheitskreis mit Periode 2π.

Raumkurven können durch ihre Graphen veranschaulicht werden, hier z.B. die Schraubenkurve

$$\gamma_1\colon \mathbb{R} \to \mathbb{R}^3\colon t \mapsto (\sin t, \cos t, t)$$

oder die Kurve

$$\gamma_2\colon \mathbb{R} \to \mathbb{R}^3\colon t \mapsto (\ln t, \sin t, t\cos t)$$



(A) γ_1 (B) γ_2

ABBILDUNG 32. Zwei Raumkurven

□

BEISPIEL 10.1.2. *Eine Funktion $f\colon D \to \mathbb{R}$ mit $D \subseteq \mathbb{R}^3$ kann zum Beispiel eine Temperaturverteilung in einem Raumbereich D beschreiben.* □

Funktionen in mehreren Variablen können auf unterschiedliche Art und Weise definiert sein: durch eine explizite Vorschrift wie z.B. $f(x, y) = x^2 + 3x^4y - 5e^y$ oder implizit durch Gleichungen wie z.B. $x^2 + y^2 + f(x, y)^2 = 1$ gegeben sein (später auch durch Differentialgleichungen).

Für die Veranschaulichung von Funktionen in mehreren Variablen gibt es verschiedene Möglichkeiten:

- Der *Graph* der Funktion $f\colon D \to \mathbb{R}^m$ mit $D \subseteq \mathbb{R}^n$ ist die Menge
$$\Gamma(f) = \{(x, f(x)) \in D \times \mathbb{R}^m \mid x \in D\}$$
Für Funktionen $f\colon D \to \mathbb{R}$ in zwei Variablen kann der Graph als Fläche im $\mathbb{R}^3$ veranschaulicht werden:
- Für die Funktion $f\colon D \to \mathbb{R}$ mit $D \subseteq \mathbb{R}^n$ kann der Verlauf der Funktion durch die *Niveaumengen*
$$N_c = \{x \in D \mid f(x) = c\}, c \in \mathbb{R}$$
veranschaulicht werden. Die Niveaumenge können wir geometrisch als Schnitt des Graphen $\Gamma_f \subseteq \mathbb{R}^n \times \mathbb{R}$ mit der Ebene $\mathbb{R}^n \times \{c\} \subseteq \mathbb{R}^n \times \mathbb{R}$ erhalten. Bekanntestes Beispiel für Niveaumengen einer Funktion sind die Höhenlinien in topographischen Landkarten oder Isobaren/Isothermen auf Wetterkarten.
- Eine andere Möglichkeit, den Funktionenverlauf von $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$ zu veranschaulichen sind die *partiellen Funktionen*
$$x_i \mapsto f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n)$$
für $(a_1, \ldots, a_n) \in \mathbb{R}^n$. Dies zeigt den Funktionenverlauf in Richtung der Variablen x_i für vorgegebene Werte $x_1 = a_1, x_2 = a_2, \ldots, x_n = a_n$. Die partiellen Funktionen können geometrisch als Schnitt des Graphen $\Gamma_f \in \mathbb{R}^n \times \mathbb{R}$ mit der durch $x_1 = a_1, x_2 = a_2, \ldots, x_n = a_n$ definierten Ebene in $\mathbb{R}^n \times \mathbb{R}$ interpretiert werden.

  Allgemeiner können wir den Graphen mit anderen Ebenen in $\mathbb{R}^n \times \mathbb{R}$, um den Verlauf der Funktion entlang verschiedener Querschnitte zu veranschaulichen.

Beispiel 10.1.3. *Betrachten wir die Funktion $f(x, y) = \cos^2(x)\cos^2(y)$. Der Graph ist in Abbildung 33 dargestellt. Der Querschnitt für $x = 0$ liefert die Funktion*

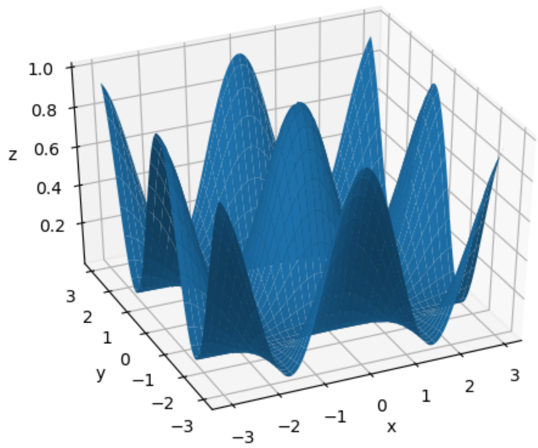


Abbildung 33. Graph der Funktion $f(x, y) = \cos^2(x)\cos^2(y)$

$f(0, y) = \cos^2(y)$, analog ist der Querschnitt für $y = 0$ die Funktion $f(x, 0) = \cos^2(x)$. Der Querschnitt für $x = \frac{\pi}{2}$ ist $f(\frac{\pi}{2}, y) = 0$. Der Querschnitt für $x = y$ (Schnitt mit der Ebene $x - y = 0$ in $\mathbb{R}^3$) ist $f(x, x) = \cos^4(x)$. □

BEISPIEL 10.1.4. *Betrachten wir die Funktion* $f(x,y) = x^2 + y^2$. *Der Graph ist in Abbildung 34 dargestellt. Die Niveaukurven* N_c *für* $c \in \mathbb{R}$ *sind Kurven der Form*

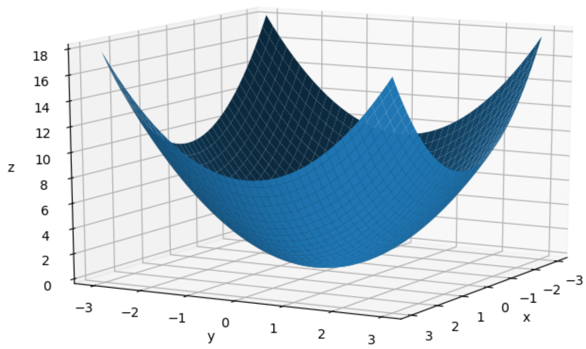


ABBILDUNG 34. Graph der Funktion $f(x,y) = x^2 + y^2$

$x^2 + y^2 = c$. *Für* $c < 0$ *gibt es keine Lösungen, d.h.* $N_c = \emptyset$, *für* $c \geq 0$ *ist* N_c *der Kreis von Radius* $\sqrt{c}$ *um den Ursprung.* □

10.2. Differenzierbare Abbildungen

Wir betrachten zuerst Ableitungen für reell-wertige Funktionen $f\colon D \to \mathbb{R}$ in mehreren Variablen, d.h. $D \subseteq \mathbb{R}^n$. In dieser Situation gibt es verschiedene Ableitungsbegriffe, da wir die Veränderung von f in verschiedene Richtungen im Definitionsbereich $D \subseteq \mathbb{R}^n$ untersuchen können. Der einfachste Begriff sind die *partiellen Ableitungen* (Ableitungen in Koordinatenrichtungen bzw. Ableitungen der partiellen Funktionen), allgemeiner gibt es *Richtungsableitungen* in beliebige Richtungen in D. Der beste Begriff von Differenzierbarkeit in diesem Kontext ist die Existenz linearer Annäherungen.

10.2.1. Ableitungen in Koordinatenrichtung: partielle Ableitungen.

DEFINITION 10.2.1. *Sei* $D \subseteq \mathbb{R}^n$ *offen und* $f\colon D \to \mathbb{R}$ *eine Funktion. Für einen festen Punkt* $p = (p_1, \ldots, p_n)$ *und einen festen Index* $i \in \{1, \ldots, n\}$ *betrachten wir die partielle Funktion* $x_i \mapsto f(p_1, \ldots, p_{i-1}, x_i, p_{i+1}, \ldots, p_n)$. *Existiert die Ableitung dieser Funktion an der Stelle* $x_i = p_i$, *dann heißt diese* partielle Ableitung von f nach x_i an der Stelle p.

Notation:

$$\frac{\partial f}{\partial x_i}(p), \qquad \left.\frac{\partial f(x)}{\partial x_i}\right|_{x=p}, \qquad f_{x_i}(p), \qquad \partial_{x_i} f(p).$$

Wenn alle partiellen Ableitungen $\frac{\partial f}{\partial x_i}(p)$, $1 \leq i \leq n$ *existieren, heißt* f partiell differenzierbar.

BEMERKUNG 10.2.2. *Nach Definition ist*

$$\frac{\partial f}{\partial x_i}(p) := \lim_{t \to 0} \frac{f(p_1, \ldots, p_{i-1}, p_i + t, p_{i+1}, \ldots, p_n) - f(p)}{t}$$

DEFINITION 10.2.3. *Sei* $D \subseteq \mathbb{R}^n$ *offen und* $f\colon D \to \mathbb{R}$ *eine Funktion, die im Punkt* $p \in D$ *partiell differenzierbar ist. Der Vektor*

$$\operatorname{grad} f(p) = \left(\frac{\partial f}{\partial x_1}(p), \ldots, \frac{\partial f}{\partial x_n}(p)\right)$$

heißt Gradient von f an der Stelle p. *Als Notation wird auch* ∇f *verwendet.*

Beispiel 10.2.4. *Für* $f(x,y) = x^3 \sin y + \mathrm{e}^x y^2$ *ist*

$$\nabla f(x,y) = \left(3x^2 \sin y + \mathrm{e}^x y^2, x^3 \cos y + \mathrm{e}^x \cdot 2y\right)$$

□

Die partiellen Ableitungen $\partial_{x_i} f$ können als Funktionen selbst wieder partiell differenzierbar sein (oder auch nicht). In diesem Fall können höhere partielle Ableitungen induktiv gebildet werden:

$$\partial_{x_i}\partial_{x_j} f = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right).$$

Beispiel 10.2.5. *Für* $f(x,y) = x^3 \sin y + \mathrm{e}^x y^2$ *ist*

$$\begin{aligned}
\partial_x\partial_x f(x,y) &= 6x \sin y + \mathrm{e}^x y^2 \\
\partial_x\partial_y f(x,y) &= 3x^2 \cos y + 2y\mathrm{e}^x \\
\partial_y\partial_x f(x,y) &= 3x^2 \cos y + 2y\mathrm{e}^x \\
\partial_y\partial_y f(x,y) &= -x^3 \sin y + 2\mathrm{e}^x
\end{aligned}$$

□

Satz 10.2.6 (Satz von Schwarz). *Sei* $D \subseteq \mathbb{R}^n$ *offen und die Funktion* $f\colon D \to \mathbb{R}$ *zweimal stetig partiell differenzierbar. Dann gilt für alle* i, j

$$\partial_{x_i}\partial_{x_j} f = \partial_{x_j}\partial_{x_i} f.$$

10.2.2. Richtungsableitungen.

Definition 10.2.7. *Sei* $D \subseteq \mathbb{R}^n$ *offen und* $f\colon D \to \mathbb{R}$ *eine Funktion. Für einen Vektor* $v \in \mathbb{R}^n$ *heißt der Grenzwert*

$$\partial_v f(p) := \lim_{t\to 0} \frac{f(p + t\cdot v) - f(p)}{t}$$

(wenn er existiert) Richtungsableitung von f am Punkt p in Richtung v.

Bemerkung 10.2.8. *Die partiellen Ableitungen* ∂_{x_i} *sind ein Spezialfall der Richtungsableitung für die Einheitsvektoren* $v = \mathrm{e}_i$. *Die Richtungsableitung* $\partial_v f$ *beschreibt den Anstieg der Funktion* f *in Richtung des Vektors* v.

Beispiel 10.2.9. *Funktion, Graph, Schnitt mit Ebene, Anstieg in Richtung, alternative Veranschaulichung mit Niveaulinien* □

Bemerkung 10.2.10. *Bemerkung zu Vektoren mit gleicher Richtung und unterschiedlicher Länge*

10.2.3. Differenzierbarkeit und Tangentialebene.

Definition 10.2.11. *Sei* $D \subseteq \mathbb{R}^n$ *offen. Eine Funktion* $f\colon D \to \mathbb{R}$ *heißt* (total) differenzierbar im Punkt $p \in D$, *wenn es eine lineare Funktion* $L\colon \mathbb{R}^n \to \mathbb{R}$ *gibt, so dass gilt*

$$\lim_{v\to 0} \frac{f(p+v) - f(p) - L(v)}{\|v\|} = 0.$$

Bemerkung 10.2.12. *(1) Hier spielt die Wahl der Norm auf* $\mathbb{R}^n$ *keine Rolle, da alle Normen auf* $\mathbb{R}^n$ *äquivalent sind, s. Satz 9.4.18.*

(2) *Die Differenzierbarkeit in der Definition ist äquivalent zur Existenz einer* linearen Annäherung*, d.h. dass sich f in einer Umgebung von p als*

$$f(x) = f(p) + L(x-p) + \phi(x-p)$$

schreiben lässt, wobei $\lim_{v\to 0} \frac{\phi(v)}{\|v\|} = 0$ ist. Bis auf einen kontrollierbaren Fehlerterm ϕ können wir f also durch die lineare Funktion $f(p)+L(x-p)$ annähern. Damit verallgemeinern wir die Beschreibung $f(x) = f(p) + f'(p)(x-p) + R(x-p)$, s. Analysis 1, Taylorentwicklung.

Proposition 10.2.13. *Sei $D \subseteq \mathbb{R}^n$ offen und $f\colon D \to \mathbb{R}$ eine im Punkt $p \in D$ differenzierbare Funktion. Dann ist f im Punkt p stetig und partiell differenzierbar. Die lineare Annäherung $L\colon \mathbb{R}^n \to \mathbb{R}$ an f ist gegeben durch Skalarprodukt mit dem Gradienten*

$$L\colon \mathbb{R}^n \to \mathbb{R}\colon v \mapsto \nabla f(p) \cdot v.$$

Beweis. Die Differenzierbarkeit bedeutet, dass wir die Funktion f linear annähern können: $f(x) = f(p) + L(x-p) + \phi(x-p)$ mit $\lim_{v\to 0} \frac{\phi(v)}{\|v\|} = 0$. Mit $\lim_{v\to 0} L(v) = 0$ folgt daraus

$$\lim_{v\to 0} f(p+v) = f(p) + \lim_{v\to 0} L(v) + \lim_{v\to 0} \phi(v) = f(p),$$

also ist f stetig in p. Die lineare Abbildung L kann in Koordinaten als

$$L(x_1,\ldots,x_n) = \sum_{i=1}^n \lambda_i x_i$$

geschrieben werden, d.h. für die lineare Annäherung von f

$$f(x) = f(p) + \sum_{i=1}^n \lambda_i(x_i - p_i) + \phi(x-p).$$

Damit ist dann $f(p + t\cdot \mathrm{e}_i) = f(p) + t\lambda_i + \phi(x-p)$ und es folgt

$$\partial_{x_i} f(p) = \lim_{t\to 0} \frac{f(p+t\mathrm{e}_i) - f(p)}{t} = \lambda_i + \lim_{t\to 0} \frac{\phi(t\mathrm{e}_i)}{t} = \lambda_i.$$

□

Bemerkung 10.2.14. (1) *Das bedeutet insbesondere, dass die lineare Abbildung L bzw. die lineare Annäherung an f eindeutig ist, wenn sie existiert. Die lineare Funktion $L\colon \mathbb{R}^n \to \mathbb{R}$ heißt* (totales) Differential *der Funktion f im Punkt p und mit $\mathrm{d}f$ bezeichnet.*

(2) *Die affin-lineare Funktion*

$$T_f(x,p) = f(p) + \nabla f(p)\cdot(x-p) = f(p) + \sum_{i=1}^n \partial_{x_i} f(p)\cdot(x_i - p_i)$$

heißt lineare Approximation von f im Punkt p. *Der Graph der Funktion T_f in $\mathbb{R}^{n+1}$ heißt* Tangentialebene *an den Graphen von f im Punkt $(p, f(p))$.*

(3) *Wenn f im Punkt $p \in D$ total differenzierbar ist, dann existieren alle Richtungsableitungen von f im Punkt p und es gilt*

$$\partial_v f(p) = \nabla f(p)\cdot v = \sum_{i=1}^n v_i \frac{\partial f}{\partial x_i}(p).$$

Der Winkel ϕ zwischen Gradient $\nabla f(p)$ und einem beliebigen Vektor v lässt sich mit der Cosinusformel berechnen

$$\cos(\phi) = \frac{v\cdot\nabla f(p)}{\|\nabla f(p)\|\cdot\|v\|}.$$

Für Vektoren v der Länge 1 ist dann $\partial_v f(p) = \|\nabla f(p)\| \cdot \cos(\phi)$. Dieser Ausdruck wird maximal für $\phi = 0$ (also wenn v parallel zu $\nabla f(p)$), und das bedeutet

$$\|\nabla f(p)\| = \max\left\{\partial_v f(p) \mid v \in \mathbb{R}^n, \|v\| = 1\right\},$$

d.h. der Gradient zeigt in Richtung des größten Anstiegs. Die Norm des Gradienten ist der maximale Wert einer Richtungsableitung (in Richtung eines normierten Vektors).

Proposition 10.2.15. *Sei $D \subseteq \mathbb{R}^n$ offen und $f\colon D \to \mathbb{R}$ eine Abbildung. Wenn f im Punkt $p \in D$ stetig partiell differenzierbar ist, dann ist f in p total differenzierbar.*

Beweis. Da D offen ist, gibt es ein $\delta > 0$, so dass $\mathrm{B}(p, \delta) \subseteq D$. Sei $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ ein Vektor mit $\|v\| < \delta$. Für $k = 0, \ldots, n$ definieren wir Punkte $x^{(k)} = p + \sum_{i=1}^k v_i \mathrm{e}_i$. Es gilt $x^{(0)} = p$ und $x^{(n)} = p + v$, und die Punkte $x^{(k-1)}$ und $x^{(k)}$ unterscheiden sich nur in der k-ten Koordinate. Nach dem Mittelwertsatz aus der Analysis 1 existieren dann $\xi_k \in [0, 1]$ mit

$$f(x^{(k)}) - f(x^{(k-1)}) = \partial_k f(y^{(k)}) v_k$$

mit $y^{(k)} = x^{(k-1)} + \xi_k v_k \mathrm{e}_k$. Daraus folgt

$$f(p + v) - f(p) = \sum_{k=1}^n \partial_k f(y^{(k)}) v_k.$$

Als lineare Annäherung an f wählen wir die lineare Funktion $L(v) = \sum_{k=1}^n \partial_k f(p) v_k$ und können dann schreiben

$$f(p + v) = f(p) + L(v) + \phi(v), \qquad \phi(v) = \sum_{k=1}^n \Big(\partial_k f(y^{(k)}) - \partial_k f(p)\Big) v_k$$

Für $v \to 0$ geht $y^{(k)}$ gegen p, und da die partiellen Ableitungen $\partial_k f$ alle stetig sind, gilt $\lim_{v\to 0} \partial_k f(y^{(k)}) = \partial_k f(p)$, also folgt

$$\lim_{v\to 0} \frac{\phi(v)}{\|v\|} = 0,$$

und damit ist f total differenzierbar. □

Bemerkung 10.2.16. *Wir haben damit also die folgenden Implikationen:* stetig partiell differenzierbar⇒ total differenzierbar⇒partiell differenzierbar. *Die Umkehrungen gelten im Allgemeinen nicht.*

10.3. Ableitungen vektorwertiger Funktionen, Kettenregel und Mittelwertsatz

Wir betrachten nun Ableitungen von vektorwertigen Funktionen $f\colon D \to \mathbb{R}^m$, $D \subseteq \mathbb{R}^n$. Solche Funktionen können wir in Komponenten zerlegen: $f(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n))$. Dann können wir über partielle Ableitungen $\partial_{x_j} f_i$ der Komponentenfunktionen f_i reden.

Definition 10.3.1. *Sei $D \subseteq \mathbb{R}^n$ offen. Eine Funktion $f\colon D \to \mathbb{R}^m$ heißt im Punkt $p \in D$* total differenzierbar*, wenn es eine lineare Abbildung $L\colon \mathbb{R}^n \to \mathbb{R}^m$ gibt, so dass in einer Umgebung von p gilt*

$$f(p + v) = f(p) + L(v) + \phi(v),$$

wobei $\phi\colon U \to \mathbb{R}^m$ eine in einer Umgebung $U \subseteq \mathbb{R}^n$ von 0 definierte Funktion mit $\lim_{v\to 0} \frac{\phi(v)}{\|v\|} = 0$ ist.

Bemerkung 10.3.2. *(1) Die Funktion $f\colon D \to \mathbb{R}^m$ ist total differenzierbar in p genau dann, wenn die Komponenten $f_i\colon D \to \mathbb{R}$ in p total differenzierbar sind.*

(2) Wir können $L\colon \mathbb{R}^n \to \mathbb{R}^m$ bezüglich der Standardbasen durch eine Matrix $A = (a_{ij})$ darstellen. Dann ist

$$L\colon v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \mapsto L(v) = A \cdot v = \begin{pmatrix} \sum_{i=1}^n a_{1i} v_i \\ \vdots \\ \sum_{i=1}^n a_{mi} v_i \end{pmatrix}$$

Die lineare Funktion L heißt Differential $\mathrm{d}f(p)$, *die darstellende Matrix heißt* Jacobi-Matrix $J_f(p) = \left(\partial_{x_j} f_i\right)_{i,j}$.

(3) Die Funktion heißt stetig differenzierbar, wenn alle partiellen Ableitungen $\partial_{x_j} f_i$ stetig sind. Für $D \subseteq \mathbb{R}^n$ offen wird die Menge der k-mal stetig differenzierbaren Funktionen $f\colon U \to \mathbb{R}^m$ mit $C^k(D, \mathbb{R}^m)$ bezeichnet.

Proposition 10.3.3 (Rechenregeln). *(1) Das Differential ist linear:*

$$\mathrm{d}(\lambda f + \mu g) = \lambda \mathrm{d}f + \mu \mathrm{d}g.$$

(2) Die Produktregel ist

$$\mathrm{d}(fg) = f\mathrm{d}g + g\mathrm{d}f.$$

Satz 10.3.4 (Kettenregel). *Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offene Mengen und $f\colon U \to \mathbb{R}^m$ und $g\colon V \to \mathbb{R}^k$ Abbildungen mit $f(U) \subseteq V$. Wenn die Funktion f im Punkt $p \in U$ differenzierbar ist und die Funktion g im Punkt $f(p)$ differenzierbar ist, dann ist die zusammengesetzte Abbildung $g \circ f\colon U \to \mathbb{R}^k$ im Punkt p differenzierbar, und es gilt*

$$\mathrm{d}(g \circ f)(p) = (\mathrm{d}g)(f(p)) \circ \mathrm{d}f(p)$$

Bemerkung 10.3.5. *Mit Jacobi-Matrizen formuliert lautet die Kettenregel*

$$J_{g\circ f}(p) = J_g(f(p)) \cdot J_f(p).$$

Beweis. Wir bezeichnen mit $A = J_f(p)$ und $B = J_g(f(p))$ die relevanten Jacobi-Matrizen. Wir müssen zeigen, dass die Jacobi-Matrix $J_{g\circ f}(p)$ das Matrizenprodukt BA ist. Die Differenzierbarkeitsvoraussetzungen für f und g bedeuten

$$f(p+v) = f(p) + A \cdot v + \phi(v), \qquad g(f(p)+w) = g(f(p)) + B \cdot w + \psi(w)$$

mit $\lim_{v\to 0} \frac{\phi(v)}{\|v\|} = \lim_{w \to 0} \frac{\psi(w)}{\|w\|} = 0$. Damit haben wir mit $w = f(p+v) - f(p) = A \cdot v + \phi(v)$ und $\chi(v) = B\phi(v) + \psi(A \cdot v + \phi(v))$

$$\begin{aligned} (g \circ f)(p+v) &= g(f(p+v)) = g(f(p) + w) \\ &= g(f(p)) + B \cdot A \cdot v + B\phi(v) + \psi(A \cdot v + \phi(v)) \\ &= (g \circ f)(p) + B \cdot A \cdot v + \chi(v). \end{aligned}$$

Für die Differenzierbarkeit von $g \circ f$ bleibt also noch zu zeigen, dass

$$\lim_{v \to 0} \frac{\chi(v)}{\|v\|} = \lim_{v \to} \frac{B\phi(v) + \psi(A \cdot v + \phi(v))}{\|v\|} = 0$$

gilt. Aus $\lim_{v\to 0} \frac{\phi(v)}{\|v\|} = 0$ folgt $\lim_{v \to 0} \frac{B \cdot \phi(v)}{\|v\|} = 0$ (Linearität von Multiplikation mit B), außerdem auch die Existenz von $\delta > 0$, so dass $\|\phi(v)\| \leq \|v\|$ für alle v mit $\|v\| \leq \delta$. Mit $\psi_1(w) = \frac{\psi(w)}{\|w\|}$ haben wir $\lim_{w\to 0} \psi_1(w) = 0$, und das bedeutet

$$\|\psi(A \cdot v + \phi(v))\| \leq (\|A\| + 1)\|v\| \cdot \|\psi_1(A \cdot v + \phi(v))\|$$

(dabei haben wir $\|A \cdot v + \phi(v)\| \leq \|A\| \cdot \|v\| + \|A\| \cdot \|\phi(v)\| \leq (\|A\| + 1)\|v\|$ benutzt). Es folgt $\lim_{v\to} \frac{\psi(A\cdot v + \phi(v))}{\|v\|} = 0$ und wir haben die Differenzierbarkeit von $g \circ f$ gezeigt. $\square$

Beispiel 10.3.6. *(1) Wir können die Richtungsableitung von $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$ am Punkt p in Richtung v als Spezialfall der Kettenregel für $f \circ \gamma$ mit der Kurve $\gamma\colon \mathbb{R} \to \mathbb{R}^n\colon t \mapsto p + tv$ auffassen.*
(2) Koordinatentransformationen sind typische Anwendungsfälle der Kettenregel.
(3) Backpropagation als Lernalgorithmus für neuronale Netzwerke ist ein Beispiel für eine praxisrelevante direkte Anwendung der Kettenregel.

□

Satz 10.3.7 (Mittelwertsatz). *Sei $D \subseteq \mathbb{R}^n$ offen und $f = (f_1, \ldots, f_m)\colon D \to \mathbb{R}^m$ differenzierbar. Seien $a, b \in D$ Punkte, so dass die Verbindungsstrecke $S(a,b) = \{a + t(b-a) \in \mathbb{R}^n \mid t \in [0,1]\}$ in D enthalten ist. Für die Matrix $M = (M_{ij})$ mit den Einträgen*

$$M_{ij} = \int_0^1 \partial_{x_j} f_i(a + t(b-a))\mathrm{d}t$$

gilt dann

$$f(b) - f(a) = M \cdot (b-a).$$

Beweis. Wir betrachten den Weg $\gamma\colon [0,1] \to D\colon t \mapsto a + t(b-a)$ und die zusammengesetzten Funktionen

$$g_i\colon [0,1] \to \mathbb{R}\colon t \mapsto g_i(t) = f_i(\gamma(t)) = f_i(a + t(b-a)).$$

Dann folgt aus der Kettenregel

$$\begin{aligned}
f_i(b) - f_i(a) &= g_i(1) - g_i(0) \\
&= \int_0^1 g_i'(t)\mathrm{d}t \\
&= \int_0^1 \left(\sum_{j=1}^n \partial_{x_j} f_i(a + t(b-a))(b_j - a_j) \right) \mathrm{d}t \\
&= \sum_{j=1}^n \left(\int_0^1 \partial_{x_j} f_i(a + t(b-a))\mathrm{d}t \right) (b_j - a_j).
\end{aligned}$$

Das ist die Behauptung. □

Bemerkung 10.3.8. *Der Spezialfall einer reellwertigen Funktion $f\colon D \to \mathbb{R}$ kann wie folgt formuliert werden:*

Sei $D \subseteq \mathbb{R}^n$ offen und $f\colon D \to \mathbb{R}$ differenzierbar. Für $a, b \in D$, so dass die Verbindungsstrecke $S(a,b) = \{a + t(b-a) \in \mathbb{R}^n \mid t \in [0,1]\}$ in D enthalten ist, existiert ein Punkt $t \in [0,1]$ mit

$$f(b) - f(a) = \nabla f(a + t(b-a)) \cdot (b-a).$$

Auch für Funktionen in mehreren Variablen gilt, dass die Beschränktheit der Ableitung Lipschitz-Stetigkeit bzw. gleichmäßige Stetigkeit impliziert. Dafür müssen wir aber noch genauer formulieren, was Beschränktheit der Ableitungen bedeuten soll.

Definition 10.3.9. *Sei $L\colon \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung mit darstellender $m \times n$-Matrix $M = (M_{ij})$. Dann definieren wir die Norm der linearen Abbildung L bzw. der Matrix M durch $\|L\| = \|M\| = \max\{\|M \cdot v\| \mid \|v\| = 1\}$.*

Korollar 10.3.10 (Schrankensatz). *Mit der Notation aus Satz 10.3.7 gilt*

$$\|f(b) - f(a)\| \leq \|M\| \cdot \|b - a\|.$$

Insbesondere ist f Lipschitz-stetig. Zusätzlich gilt

$$\|M\| \leq \sup\{\|\partial_i f(p+tv)\| \mid 1 \leq i \leq n, t \in [0,1]\}.$$

Beweis. Aus Satz 10.3.7 folgt

$$f(b) - f(a) = M \cdot (b-a).$$

Die Aussage folgt aus der Übungsaufgabe.

Die zweite Abschätzung folgt aus

$$\left\| \int_0^1 \partial_{x_j} f_i(a+t(b-a))\mathrm{d}t \right\| \leq \int_0^1 \left\|\partial_{x_j} f_i(a+t(b-a))\right\| \mathrm{d}t$$

□

Übungsaufgabe 10.3.11. *Sei $L\colon \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung mit darstellender Matrix $M = (M_{ij})$ und $v \in \mathbb{R}^n$ ein Vektor. Dann gilt*

$$\|L(v)\| = \|M \cdot v\| \leq \|M\| \cdot \|v\|.$$

10.4. Taylorentwicklung und Hesse-Matrix

Wir betrachten wieder reell-wertige Funktionen $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$. Wie für Funktionen in einer Variablen geht es bei der Taylor-Entwicklung darum, eine Funktion in mehreren Variablen durch Polynome anzunähern. Dafür benötigen wir etwas Notation, um mit den verschiedenen Potenzen der verschiedenen Variablen umzugehen. Für ein n-Tupel $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ bezeichnen wir

$$|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n, \qquad \alpha! = \alpha_1!\alpha_2! \cdots \alpha_n!$$

Für einen Punkt $x = (x_1, \dots, x_n)$ bezeichnen wir $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, damit hat ein Polynom in n Variablen dann die Form

$$f(x) = \sum_{\alpha \in \mathbb{N}^n} \lambda_\alpha x^\alpha.$$

Für ein Indextupel $\alpha \in \mathbb{N}^n$ und eine $|\alpha|$-mal stetig differenzierbare Funktion $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$ definieren wir

$$\mathrm{D}^\alpha f = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n} f,$$

wobei $\partial_{x_i}^{\alpha_i} = \underbrace{\partial_{x_i}\partial_{x_i} \cdots \partial_{x_i}}_{\alpha_i\text{-mal}}$ ist. Hier benutzen wir den Satz von Schwarz, damit die Reihenfolge der partiellen Ableitungen in dieser Definition keine Rolle spielt.

Satz 10.4.1. *Sei $D \subseteq \mathbb{R}^n$ offen und $f\colon D \to \mathbb{R}$ eine k-mal stetig differenzierbare Funktion. Für einen Punkt $p \in D$ und einen Vektor $v \in \mathbb{R}^n$, so dass die Verbindungsstrecke $S(p, p+v) = \{p+tv \mid t \in [0,1]\}$ in D liegt, ist die Funktion*

$$g\colon [0,1] \to \mathbb{R}\colon t \mapsto f(p+tv)$$

k-mal stetig differenzierbar und es gilt

$$\frac{\mathrm{d}^k g}{\mathrm{d}t^k}(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!}(\mathrm{D}^\alpha f)(p+tv)v^\alpha.$$

Beweis. Zuerst zeigen wir durch Induktion über k, dass gilt

$$\frac{\mathrm{d}^k g}{\mathrm{d}t^k}(t) = \sum_{i_1,\dots,i_k=1}^{n} \partial_{i_k} \cdots \partial_{i_1} f(p+tv) v_{i_1} \cdots v_{i_k}$$

Der Induktionsanfang $k=1$ folgt aus der Kettenregel

$$\frac{\mathrm{d}g}{\mathrm{d}t}(t) = \frac{\mathrm{d}}{\mathrm{d}t} f(p_1 + tv_1, \dots, p_n + tv_n) = \sum_{i=1}^n \partial_i f(p+tv)v_i.$$

Für den Induktionsschritt benutzen wir auch wieder die Kettenregel:

$$\begin{aligned}\frac{\mathrm{d}^k g}{\mathrm{d}t^k}(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i_1,\ldots,i_{k-1}=1}^{n} \partial_{i_{k-1}}\cdots\partial_{i_1} f(p+tv)v_{i_1}\cdots v_{i_{k-1}}\right)\\ &= \sum_{j=1}^{n}\partial_j\left(\sum_{i_1,\ldots,i_{k-1}=1}^{n} \partial_{i_{k-1}}\cdots\partial_{i_1} f(p+tv)v_{i_1}\cdots v_{i_{k-1}}\right)v_j\\ &= \sum_{i_1,\ldots,i_k=1}^{n} \partial_{i_k}\cdots\partial_{i_1} f(p+tv)v_{i_1}\cdots v_{i_k}\end{aligned}$$

In dieser Formel haben wir die Reihenfolge der Indizes genau vorgegeben, unter $(i_1,\ldots,i_k)$ können die Indizes $1,\ldots,n$ natürlich mehrfach vorkommen. Jetzt können wir mit dem Satz von Schwarz die Ableitungsreihenfolge vertauschen. Die Anzahl der Tupel $(i_1,\ldots,i_k)$ mit $1 \le i_j \le n$, in denen der Index ν genau α_ν-mal vorkommt (d.h. insbesondere, dass $k = |\alpha|$), ist

$$\frac{k!}{\alpha_1!\cdots\alpha_n!}$$

Daraus folgt die Behauptung. □

Satz 10.4.2 (Taylor-Entwicklung). *Sei $D \subseteq \mathbb{R}^n$ offen, und sei $p \in D$ ein Punkt und $v \in \mathbb{R}^n$ ein Vektor, so dass die Verbindungsstrecke $S(p,p+v) = \{p+tv \mid t \in [0,1]\}$ in D liegt. Für eine $k+1$-mal stetig differenzierbare Funktion $f\colon D \to \mathbb{R}$ existiert ein $\theta \in [0,1]$, so dass*

$$f(p+v) = \sum_{|\alpha|\le k}\frac{\mathrm{D}^\alpha f(p)}{\alpha!}v^\alpha + \sum_{|\alpha|=k+1}\frac{\mathrm{D}^\alpha f(p+\theta v)}{\alpha!}v^\alpha$$

Beweis. Wir betrachten $g\colon [0,1] \to \mathbb{R}\colon t \mapsto f(p+tv)$. Nach Satz 10.4.1 ist g eine $k+1$-mal stetig differenzierbare Funktion, und nach Taylorentwicklung in einer Variablen existiert ein $\theta \in [0,1]$ mit

$$g(1) = \sum_{m=0}^{k}\frac{g^{(m)}(0)}{m!} + \frac{g^{(k+1)}(\theta)}{(k+1)!}.$$

Aus Satz 10.4.1 bekommen wir

$$\frac{g^{(m)}(0)}{m!} = \sum_{|\alpha|=m}\frac{\mathrm{D}^\alpha f(p)}{\alpha!}v^\alpha, \qquad \frac{g^{(k+1)}(\theta)}{(k+1)!} = \sum_{|\alpha|=k+1}\frac{\mathrm{D}^\alpha f(p+\theta v)}{\alpha!}v^\alpha,$$

und damit die Behauptung. □

Korollar 10.4.3. *Für das Restglied $R_{k+1}(p,v) = \sum_{|\alpha|=k+1}\frac{\mathrm{D}^\alpha f(p+\theta v)}{\alpha!}v^\alpha$ aus Satz 10.4.2 gilt*

$$\lim_{v\to 0}\frac{R_{k+1}(p,v)}{\|v\|^k} = \lim_{v\to 0}\frac{f(p+v)-T_k(p,v)}{\|v\|^p} = 0.$$

Bemerkung 10.4.4. *Das Taylorpolynom vom Grad k ist also die beste Annäherung an eine $k+1$-mal stetig differenzierbare Funktion, die mit einem Polynom vom Grad k möglich ist. Es zerlegt sich in homogene Summanden*

$$P_m(v) = \sum_{|\alpha|=m}\frac{\mathrm{D}^\alpha f(p)}{\alpha!}v^\alpha.$$

Im Grad 0 haben wir $P_0(v) = \frac{\mathrm{D}^0 f(p)}{0!}v^0 = f(p)$, das ist also der Funktionswert im Entwicklungspunkt p.

Für $m = 1$ haben wir mit $|\alpha| = 1$ für $\alpha = i$ nur die einfachen partiellen Ableitungen $\mathrm{D}^\alpha = \partial_i$. Damit ist

$$P_1(v) = \sum_{i=1}^{n} \partial_i f(p) v_i = \nabla f(p) \cdot v.$$

Das Taylorpolynom vom Grad 1 ist also die lineare Approximation der Funktion:

$$f(p+v) = f(p) + \nabla f(p) \cdot v.$$

Für $m = 2$, d.h. $|\alpha| = 2$, sind D^α die zweifachen partiellen Ableitungen von f, und

$$P_2(v) = \frac{1}{2} \sum_{i=1}^{n} \partial_{x_i}^2 f(p) v_i^2 + \sum_{i<j} \partial_{x_i} \partial_{x_j} f(p) v_i v_j = \frac{1}{2} \sum_{i,j=1}^{n} \partial_{x_i} \partial_{x_j} f(p) v_i v_j.$$

DEFINITION 10.4.5. *Sei $D \subseteq \mathbb{R}^n$ offen und $f \colon D \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Die* Hesse-Matrix *von f im Punkt $p \in D$ ist die symmetrische Matrix:*

$$H_f(p) = \big(\partial_{x_i} \partial_{x_j} f(p)\big)_{i,j}$$

BEMERKUNG 10.4.6. *Sei $D \subseteq \mathbb{R}^n$ offen und $f \colon D \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Das Taylorpolynom zweiten Grades für f mit Entwicklungspunkt p ist*

$$T_2(p+v) = f(p) + \nabla f(p) \cdot v + \frac{1}{2} v^{\mathrm{t}} \cdot H_f(p) \cdot v$$

10.5. Lokale Extrema

DEFINITION 10.5.1. *Eine Funktion $f \colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$ hat im Punkt $p \in D$ ein* lokales Maximum *(bzw.* Minimum*), wenn eine Umgebung $p \in U \subseteq D$ existiert, so dass für alle $x \in U$ gilt $f(x) \leq f(p)$ (bzw. $f(x) \geq f(p)$).*

BEMERKUNG 10.5.2. *Wenn es eine Umgebung U gibt, so dass für alle Punkte $x \in U \setminus \{p\}$ gilt $f(x) < f(p)$ bzw. $f(x) > f(p)$, dann sprechen wir von* striktem *oder* isoliertem *Maximum bzw. Minimum. Maxima und Minima werden als* Extrema *bezeichnet, p ist die* Extremstelle *und $f(p)$ der* Extremwert.

Wir betrachten zuerst Extremstellen im Innern des Definitionsbereichs. Extremstellen auf dem Rand sind komplizierter (Extremstellen mit Nebenbedingungen) und werden nach der Diskussion impliziter Funktionen betrachtet.

PROPOSITION 10.5.3 (Notwendiges Kriterium für Extremstellen). *Sei $D \subseteq \mathbb{R}^n$ offen und $f \colon D \to \mathbb{R}$ differenzierbar. Wenn f in $p \in D$ ein lokales Extremum besitzt, gilt $\nabla f(p) = 0$.*

BEWEIS. Für alle Einheitsvektoren $\mathrm{e}_1, \ldots, \mathrm{e}_n$ hat die Funktion $F(t) := f(p + t\mathrm{e}_k)$ im Punkt $t = 0$ ein lokales Extremum. Die Aussage folgt dann aus dem notwendigen Extremstellenkriterium aus Analysis 1. □

BEMERKUNG 10.5.4. *Geometrisch bedeutet das, dass die Tangentialebene an den Graphen $\Gamma(f) \subseteq \mathbb{R}^n \times \mathbb{R}$ parallel zur Ebene $\mathbb{R}^n \times \{0\}$ ist (d.h. waagerecht); alternativ: das erste Taylorpolynom ist konstant. Für eine Funktion $f \colon D \to \mathbb{R}$ heißen Punkte p mit $\nabla f(p) = 0$* kritische Punkte.

Definition 10.5.5. *Sei M eine symmetrische $n \times n$-Matrix mit reellen Einträgen. Dann heißt M*

- positiv (semi-)definit, *wenn für alle Vektoren $0 \neq v \in \mathbb{R}^n$ gilt $v^{\mathrm{t}} \cdot M \cdot v > 0$ (bzw. $v^{\mathrm{t}} \cdot M \cdot v \geq 0$),*
- negativ (semi-)definit, *wenn für alle Vektoren $0 \neq v \in \mathbb{R}^n$ gilt $v^{\mathrm{t}} \cdot M \cdot v < 0$ (bzw. $v^{\mathrm{t}} \cdot M \cdot v \leq 0$),*
- indefinit, *wenn es Vektoren $v, w \in \mathbb{R}^n$ gibt, so dass $v^{\mathrm{t}} \cdot M \cdot v > 0$ und $w^{\mathrm{t}} \cdot M \cdot w < 0$.*

Bemerkung 10.5.6. *Reelle symmetrische Matrizen sind diagonalisierbar (s. Vorlesung Lineare Algebra 2) und alle Eigenwerte sind reell. Das bedeutet, dass es für eine reelle symmetrische $n \times n$-Matrix M eine Basis $v_1, \ldots, v_n$ von $\mathbb{R}^n$ und reelle Zahlen $\lambda_1, \ldots, \lambda_n$ gibt, so dass $M \cdot v_i = \lambda_i v_i$ gilt. Die Matrix M ist*

- *positiv (semi-)definit, wenn alle $\lambda_i > 0$ (bzw. $\lambda_i \geq 0$) sind,*
- *negativ (semi-)definit, wenn alle $\lambda_i < 0$ (bzw. $\lambda_i \leq 0$) sind,*
- *indefinit, wenn es positive und negative λ_i gibt.*

Ein einfacheres Kriterium für positive Definitheit von Matrizen ist das folgende:

Proposition 10.5.7 (Sylvester-Kriterium/Hurwitz-Kriterium). *Sei $M = (m_{ij})_{i,j}$ eine reelle symmetrische $n \times n$-Matrix. Dann ist die Matrix M positiv definit, wenn für alle $1 \leq k \leq n$ gilt*

$$\det \begin{pmatrix} m_{11} & \cdots & m_{1k} \\ \vdots & \ddots & \vdots \\ m_{k1} & \cdots & m_{kk} \end{pmatrix} > 0.$$

Bemerkung 10.5.8. *Damit ist z.B. eine reelle symmetrische 2×2-Matrix $M = (m_{ij})_{i,j}$ positiv definit, wenn $m_{11} > 0$ und $\det M = m_{11}m_{22} - m_{12}m_{21} > 0$ ist.*

Proposition 10.5.9. *Sei $D \subseteq \mathbb{R}^n$ offen, $f \colon D \to \mathbb{R}$ zweimal stetig differenzierbar und sei $p \in D$ ein kritischer Punkt.*

(1) Wenn $H_f(p)$ positiv definit ist, hat f in p ein lokales Minimum.
(2) Wenn $H_f(p)$ negativ definit ist, hat f in p ein lokales Maximum.
(3) Wenn $H_f(p)$ indefinit ist, hat f in p kein lokales Extremum.

Bemerkung 10.5.10. *Sei $D \subseteq \mathbb{R}^n$ offen, $f \colon D \to \mathbb{R}$ zweimal stetig differenzierbar und $p \in D$ ein kritischer Punkt. Die Diagonalisierbarkeit der Hesse-Matrix für f bedeutet, dass in einer Umgebung von p und in geeigneten Koordinaten die quadratische Annäherung (Taylorpolynom Grad 2) die folgende Form hat:*

$$T_2(x,p) = f(c) + \sum_{i=1}^{n} \lambda_i (x_i - p_i)^2.$$

Die λ_i sind dabei die Eigenwerte der Hesse-Matrix $H_f(p)$. Die Entwicklung der Funktionswerte in der Umgebung von p kann dann (nach Abschätzung des Restterms) an den λ_i abgelesen werden. Sind alle $\lambda_i > 0$ (bzw. $\lambda_i < 0$), werden die Funktionswerte in alle Richtungen größer (bzw. kleiner), wir haben also ein lokales Minimum (bzw. Maximum). Gibt es sowohl positive als auch negative λ_i, gibt es Richtungen, in die die Funktionswerte größer werden und andere, in die die Funktionswerte kleiner werden, wir haben also kein lokales Extremum.

10.6. Implizite Funktionen und Umkehrsatz

Definition 10.6.1. *Sei (X, d) ein metrischer Raum. Eine Abbildung $f \colon X \to X$ heißt* Kontraktion, *wenn es eine Zahl $\lambda < 1$ gibt, so dass für alle $x, y \in X$ gilt*

$$d(f(x), f(y)) \leq \lambda d(x, y).$$

BEISPIEL 10.6.2. *Lipschitz-stetige Abbildungen $f\colon \mathbb{R}^n \to \mathbb{R}^n$ mit Lipschitz-Konstante $L < 1$ sind Kontraktionen. Kontraktionen sind Lipschitz-stetig und damit stetig.* □

SATZ 10.6.3 (BANACHscher Fixpunktsatz). *Sei $X \neq \emptyset$ ein vollständiger metrischer Raum und $f\colon X \to X$ eine Kontraktion. Dann gibt es genau einen Fixpunkt, d.h. einen Punkt $p \in X$ mit $f(p) = p$. Für jedes $x_0 \in X$ konvergiert die rekursive definierte Folge $x_{n+1} = f(x_n)$ gegen den Fixpunkt p.*

BEWEIS. Für die Eindeutigkeit seien p, q zwei Fixpunkte. Dann gilt

$$d(p,q) = d(f(p), f(q)) \leq \lambda d(p,q),$$

daraus folgt aber $d(p,q) = 0$.

Für die Existenz sei $x_0 \in X$ beliebig, wir betrachten die rekursiv definierte Folge $(x_n)_{n\in\mathbb{N}}$ mit dem gegebenen x_0 und $x_{n+1} = f(x_n)$ für $n \geq 0$. Dann gilt

$$d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leq \lambda d(x_n, x_{n-1}) \leq \lambda^2 d(x_{n-1}, x_{n-2}) \leq \lambda^n d(x_1, x_0).$$

Daraus folgt mit der Dreiecksungleichung

$$\begin{aligned} d(x_m, x_n) &\leq d(x_m, x_{m-1}) + d(x_{m-1}, x_{m-2}) + \cdots + d(x_{n+1}, x_n) \\ &\leq (\lambda^{m-1} + \lambda^{m-2} + \cdots + \lambda^n) d(x_1, x_0) \\ &\leq d(x_1, x_0) \sum_{k=n}^{\infty} \lambda^k \end{aligned}$$

Da $0 \leq \lambda < 1$ haben wir $\sum_{k=0}^{\infty} \lambda^k = \frac{1}{1-\lambda}$, und damit ist $\lim_{n\to\infty} \sum_{k=n}^{\infty} \lambda^k = 0$. Damit geht $d(x_m, x_n)$ für $m > n$ und $n \to \infty$ gegen 0, d.h. (x_n) ist eine Cauchy-Folge. Da X vollständig ist, konvergiert (x_n) gegen einen Grenzwert p. Für diesen Grenzwert p gilt (mit Stetigkeit von f)

$$f(p) = f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = p.$$

□

BEMERKUNG 10.6.4. *Vergleiche auch Satz 6.4.10, Anwendungen: Newton-Verfahren zur Nullstellenbestimmung, Lösung von Integralgleichungen (durch Kontraktion auf Raum der stetigen Funktionen auf $[a,b]$ mit Supremumsnorm), Existenz und Eindeutigkeit von Lösungen von gewöhnlichen Differentialgleichungen erster Ordnung (Picard–Lindelöf).*

ÜBUNGSAUFGABE 10.6.5. *Für $1 \leq i, k \leq n$ seien reelle Zahlen b_i und c_{ik} gegeben, so dass*

$$\sum_{i,k=1}^{n} c_{ik}^2 < 1.$$

Zeigen Sie mit Hilfe des Fixpunktsatzes von Banach, dass das nichtlineare Gleichungssystem

$$x_i = b_i + \sum_{k=1}^{n} \sin(c_{ik} x_j), \quad 1 \leq i \leq n$$

genau eine Lösung besitzt.

SATZ 10.6.6 (Satz über implizite Funktionen). *Seien $U \subseteq \mathbb{R}^k$ und $V \subseteq \mathbb{R}^m$ offene Teilmengen und*

$$F\colon U \times V \to \mathbb{R}^m\colon (x,y) \mapsto F(x,y)$$

stetig differenzierbar. Sei $(p,q) \in U \times V$ ein Punkt mit $F(p,q) = 0$, so dass die Jacobi-Matrix

$$\frac{\partial F}{\partial y} = \begin{pmatrix} \partial_{y_1} F_1 & \cdots & \partial_{y_m} F_1 \\ \vdots & \ddots & \vdots \\ \partial_{y_1} F_m & \cdots & \partial_{y_m} F_m \end{pmatrix}$$

im Punkt (p,q) invertierbar ist. Dann existieren offene Umgebungen $U' \subseteq U$ von p und $V' \subseteq V$ von q und eine stetig differenzierbare Abbildung $g\colon U' \to V' \subseteq \mathbb{R}^m$ mit $g(p) = q$, so dass für alle $(x,y) \in U' \times V'$ gilt $F(x,y) = 0$ genau dann, wenn $y = g(x)$.

Beweis. Nach linearer Koordinatentransformation können wir annehmen, dass $(p,q) = (0,0)$ ist. Mit der Matrix $B = \frac{\partial F}{\partial y}(0,0)$ definieren wir

$$G(x,y) := y - B^{-1}F(x,y).$$

Mit dieser Definition ist $F(x,y) = 0$ genau dann, wenn $G(x,y) = y$, damit muss ein Fixpunktproblem gelöst werden.

Wir müssen in eine Situation kommen, in der der Fixpunktsatz 10.6.3 von Banach anwendbar ist. Die Ableitung ist

$$\frac{\partial G}{\partial y}(x,y) = \mathrm{I}_m - B^{-1}\frac{\partial F}{\partial y}(x,y),$$

wobei I_m die $m \times m$-Einheitsmatrix ist. Das bedeutet $\frac{\partial G}{\partial y}(0,0) = 0$, und da alle Komponenten der Matrix $\frac{\partial G}{\partial y}$ stetig sind, gibt es Umgebungen $W_1 \subseteq U$ und $W_2 \subseteq V$ von 0, so dass

$$\left\| \frac{\partial G}{\partial y}(x,y) \right\| \leq \frac{1}{2}$$

für alle $(x,y) \in W_1 \times W_2$ gilt. Das bedeutet (nach Taylorentwicklung) insbesondere $\|G(x,y) - G(x,z)\| \leq \frac{1}{2}\|y - z\|$ für $x \in W_1$ und $y, z \in W_2$.

Dann wählen wir $r > 0$, so dass $V' := \{y \in \mathbb{R}^m \mid \|y\| \leq r\} \subseteq W_2$. Wegen $G(0,0) = 0$ und Stetigkeit existiert eine offene Umgebung U' von 0, so dass

$$\sup_{x \in U'} \|G(x,0)\| \leq \frac{r}{2}.$$

Aus der vorigen Abschätzung für die Funktionswerte von G erhalten wir mit $z = 0$ dann für alle $x \in U'$, dass aus $\|y\| \leq r$ auch $\|G(x,y)\| \leq r$ folgt.

Wir wollen jetzt den Fixpunktsatz von Banach anwenden. Dazu betrachten wir den metrischen Raum $C_b(U', \mathbb{R}^m)$ der stetigen beschränkten Abbildungen $U' \to \mathbb{R}^m$ mit der Supremumsnorm $\|f\|_\infty := \sup\{\|f(x)\| \mid x \in U'\}$ an. Wir definieren eine Abbildung

$$\Phi\colon C_b(U', \mathbb{R}^m) \to C_b(U', \mathbb{R}^m)\colon (f\colon U' \to \mathbb{R}^m) \mapsto (\Phi(f)\colon U' \to \mathbb{R}^m\colon x \mapsto G(x, f(x)))$$

Für eine Abbildung $f\colon U' \to \mathbb{R}^m$ mit $\|f\|_\infty \leq r$ gilt nach den obigen Abschätzungen auch $\|\Phi(f)\|_\infty \leq r$. Das bedeutet, dass sich Φ zu einer Abbildung $\Phi\colon A \to A$ auf der abgeschlossenen Teilmenge

$$A = \{f \in C_b(U', \mathbb{R}^m) \mid \|f\|_\infty \leq r\} = \{f \in C_b(U', \mathbb{R}) \mid f(U') \subseteq V'\}$$

einschränken lässt. Aus den obigen Abschätzungen folgt

$$\begin{aligned} \|\Phi(f) - \Phi(f')\|_\infty &= \sup_{x \in U'} \|G(x, f(x)) - G(x, f'(x))\| \\ &\leq \frac{1}{2} \sup_{x \in U'} \|f(x) - f'(x)\| \\ &= \frac{1}{2}\|f - f'\|_\infty, \end{aligned}$$

damit ist Φ eine Kontraktion. Nach Satz 10.6.3 existiert ein Fixpunkt $g \in A \subseteq C_b(U', \mathbb{R}^m)$, d.h. eine *stetige* Abbildung $g\colon U' \to V'$ mit $G(x, g(x)) = g(x)$ für alle $x \in U'$, und das bedeutet $F(x, g(x)) = 0$ für alle $x \in U'$.

Für den Beweis der stetigen Differenzierbarkeit von $g\colon U' \to \mathbb{R}^m$ können wir nach Verkleinerung von U' annehmen, dass $\frac{\partial F}{\partial y}$ in allen Punkten $(x, g(x))$ invertierbar ist. Es reicht, die Differenzierbarkeit im Punkt $(0,0)$ zu zeigen, für die anderen Punkte ist der Beweis analog. Die Differenzierbarkeit von F in $(0,0)$ bedeutet

$$F(x,y) = Ax + By + \phi(x,y), \qquad \lim_{(x,y)\to 0} \frac{\phi(x,y)}{\|(x,y)\|} = 0,$$

wobei $A = \frac{\partial F}{\partial x}(0,0)$ und $B = \frac{\partial F}{\partial y}(0,0)$ ist. Mit $F(x, g(x)) = 0$ für $x \in U'$ folgt daraus durch Umstellen

$$g(x) = -B^{-1}Ax - B^{-1}\phi(x, g(x)),$$

und es bleibt zu zeigen, dass $\lim_{x\to 0} \frac{-B^{-1}\phi(x,g(x))}{\|x\|} = 0$. Dafür zeigt man $\|g(x)\| \le K\|x\|$, was aus der Abschätzung für $\phi(x,y)$ folgt, für Details s. Forster Analysis 2. Die Formel für die Jacobi-Matrix von g im nachfolgenden Korollar zeigt, dass die Abbildung g *stetig* differenzierbar ist. $\square$

Korollar 10.6.7. *In der Situation von Satz 10.6.6 gilt*

$$\frac{\partial g}{\partial x}(x) = -\left(\frac{\partial F}{\partial y}(x, g(x))\right)^{-1} \frac{\partial F}{\partial x}(x, g(x)).$$

Beweis. Durch partielle Ableitung der i-ten Komponente von $F(x, g(x)) = 0$ nach x_j erhalten wir mit der Kettenregel

$$\frac{\partial F_i}{\partial x_j}(x, g(x)) + \sum_{k=1}^{m} \frac{\partial F_i}{\partial y_k}(x, g(x)) \frac{\partial g_k}{\partial x_j}(x) = 0.$$

Mit Matrizenprodukt ist das zusammengefasst

$$\frac{\partial F}{\partial x}(x, g(x)) + \frac{\partial F}{\partial y}(x, g(x)) \frac{\partial g}{\partial x}(x) = 0.$$

Wenn die Matrix $\frac{\partial F}{\partial y}$ im Punkt $(x, g(x))$ invertierbar ist, folgt die Behauptung durch umstellen der Gleichung. $\square$

Satz 10.6.8 (Umkehrsatz). *Sei $D \subseteq \mathbb{R}^n$ offen und $f\colon D \to \mathbb{R}^n$ eine stetig differenzierbare Abbildung. Sei $p \in D$ und $q = f(p)$. Wenn die Jacobi-Matrix $J_f(p)$ invertierbar ist, dann gibt es eine offene Umgebung $U \subseteq D$ von p und eine offene Umgebung V von q, so dass $f\colon U \to V$ bijektiv ist und die Umkehrabbildung $f^{-1}\colon V \to U$ stetig differenzierbar mit Jacobi-Matrix $J_g(q) = J_f(p)^{-1}$ ist.*

Beweis. Wir definieren die Abbildung

$$F\colon \mathbb{R}^n \times D \to \mathbb{R}^n \colon (x,y) \mapsto F(x,y) := x - f(y).$$

Dann ist $F(q,p) = 0$ und $\frac{\partial F}{\partial y}(x,y) = -J_f(y)$. Da $J_f(p)$ invertierbar ist, folgt aus dem Satz über implizite Funktionen 10.6.6, dass es offene Umgebungen U von p und V von q und eine stetige Abbildung $g\colon V \to U$, so dass für alle $(x,y) \in V \times U$ genau dann $F(x,y) = x - f(y) = 0$ gilt, wenn $y = g(x)$ ist. Durch eventuelles Verkleinern von U und V erreichen wir, dass $f\colon U \to V$ eine Bijektion mit Umkehrung $g\colon V \to U$ ist. $\square$

Bemerkung 10.6.9. *Die Bedingung im Umkehrsatz ist (aus der Kettenregel) eine notwendige Bedingung für die Umkehrbarkeit, der Umkehrsatz sagt, dass diese Bedingung lokal auch hinreichend ist.*

Übungsaufgabe 10.6.10. *Sei* $f(x_1, x_2) = (x_1^2 - x_2^2, 2x_1x_2)$.

(1) *Zeigen Sie, dass f in jedem Punkt $(x_1, x_2) \neq (0,0)$ lokal umkehrbar ist. Ist f als Abbildung von $\mathbb{R}^2 \setminus \{(0,0)\}$ in sich global umkehrbar?*
(2) *Finden Sie eine affine Abbildung, die die lokale Umkehrung f^{-1} in der Nähe von $f(1,-1)$ approximiert.*

Definition 10.6.11. *Seien $U, V \subseteq \mathbb{R}^n$ offen. Eine bijektive stetig differenzierbare Abbildung $f\colon U \to V$ heißt* Diffeomorphismus, *wenn auch die Umkehrabbildung $f^{-1}\colon V \to U$ stetig differenzierbar ist.*

Beispiel 10.6.12. *Standardbeispiele für Diffeomorphismen sind Koordinatentransformationen.* □

10.7. Anwendung: Extrema mit Nebenbedingungen

In Abschnitt 10.5 haben wir Extrema für Funktionen auf offenen Definitionsbereichen betrachtet. Jetzt betrachten wir eine etwas andere Situation, in der die Extremstellen zusätzliche Nebenbedingungen erfüllen sollen. Sei $D \subseteq \mathbb{R}^n$ offen und seien $f\colon D \to \mathbb{R}$ und $g\colon D \to \mathbb{R}^m$ stetig differenzierbare Abbildungen mit $n > m$. Wir suchen jetzt Extremstellen x von f, die der Nebenbedingung $g(x) = 0$ genügen, die also in der Nullstellenmenge $M = \{x \in U \mid g(x) = 0\}$ liegen. Die Lagrange-Multiplikatoren beschreiben eine notwendige Bedingung, die solche lokalen Extremstellen erfüllen müssen:

Satz 10.7.1 (Multiplikatorregel von Lagrange). *Sei $D \subseteq \mathbb{R}^n$ offen, $\phi\colon D \to \mathbb{R}$ und $F\colon D \to \mathbb{R}^m$ stetig differenzierbar, und $n > m$. Wir nehmen an, dass $a \in D$ eine lokale Extremstelle von ϕ ist, $F(a) = 0$ gilt und dass die Matrix*

$$\frac{\partial F}{\partial z} = \begin{pmatrix} \partial_{z_1} F_1 & \cdots & \partial_{z_n} F_1 \\ \vdots & \ddots & \vdots \\ \partial_{z_1} F_m & \cdots & \partial_{z_n} F_m \end{pmatrix}$$

vollen Rang hat (d.h. dass die Gradienten $\nabla F_1, \ldots, \nabla F_m$ in a linear unabhängig sind). Dann existieren Zahlen $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$, so dass gilt

$$\nabla\phi(a) = \sum_{j=1}^{m} \lambda_j \nabla F_j(a).$$

Beweis. Die Matrix $\frac{\partial F}{\partial z}$ hat vollen Rang, also (nach LA-Vorlesung) m linear unabhängige Spalten. Die dazugehörigen Variablen benennen wir in $y_1, \ldots, y_m$ um, die restlichen $k = n - m$ Variablen benennen wir in $x_1, \ldots, x_k$ um. Dann ist

$$F\colon D \subseteq \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^m$$

eine Funktion, auf die der Satz 10.6.6 über implizite Funktionen angewendet werden kann: wir betrachten hierbei den Punkt $a = (p, q) \in \mathbb{R}^k \times \mathbb{R}^m$, für den $F(p, q) = 0$ gilt, und nach Konstruktion ist die Matrix

$$\frac{\partial F}{\partial y} = \begin{pmatrix} \partial_{y_1} F_1 & \cdots & \partial_{y_m} F_1 \\ \vdots & \ddots & \vdots \\ \partial_{y_1} F_m & \cdots & \partial_{y_m} F_m \end{pmatrix}$$

invertierbar. Nach Satz 10.6.6 existieren offene Umgebungen $U' \subseteq \mathbb{R}^k$ von p und $V' \subseteq \mathbb{R}^m$ von q und eine stetig differenzierbare Abbildung $g\colon U' \to V'$, so dass für alle $(x, y) \in U' \times V'$ genau dann $F(x, y) = 0$ gilt, wenn $y = g(x)$. Damit haben wir die Menge $M = \{(x, y) \in D \mid F(x, y) = 0\}$ (in der die Nebenbedingung $F(x, y) = 0$) erfüllt ist) durch die Variablen $x_1, \ldots, x_k$ parametrisiert.

Wir betrachten jetzt die Funktionen $\mathrm{id} \times g \colon U' \subseteq \mathbb{R}^k \to \mathbb{R}^k \times \mathbb{R}^m$, die durch

$$(x_1, \dots, x_k) \mapsto (x_1, \dots, x_k, g_1(x_1, \dots, x_k), \dots, g_m(x_1, \dots, x_k))$$

gegeben ist. Die Jacobi-Matrix dieser Funktion ist

$$J_{\mathrm{id}\times g} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ \partial_{x_1} g_1 & \partial_{x_2} g_1 & \cdots & \partial_{x_k} g_1 \\ \partial_{x_1} g_2 & \partial_{x_2} g_2 & \cdots & \partial_{x_k} g_2 \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} g_m & \partial_{x_2} g_m & \cdots & \partial_{x_k} g_m \end{pmatrix}$$

Das ist also eine $n \times k$-Matrix deren erste k Zeilen eine $k \times k$-Einheitsmatrix sind und deren letzte $m \times k$ Zeilen die Matrix $\frac{\partial g}{\partial x}$ sind. Nach Korollar 10.6.7 haben wir

$$\frac{\partial g}{\partial x}(x) = -\left(\frac{\partial F}{\partial y}(x, g(x))\right)^{-1} \frac{\partial F}{\partial x}(x, g(x)).$$

Nun setzen wir die Funktion $\mathrm{id} \times g$ mit der Funktion $\phi \colon \mathbb{R}^n \to \mathbb{R}$ zusammen zu einer Funktion

$$\Phi \colon U' \subseteq \mathbb{R}^k \xrightarrow{\mathrm{id}\times g} \mathbb{R}^n \xrightarrow{\phi} \mathbb{R}$$

und benutzen die Kettenregel, um die Jacobi-Matrix dieser zusammengesetzten Funktion (in diesem Fall, den Gradienten) auszurechnen:

$$\nabla\Phi(x) = J_\Phi(x) = J_\phi(x, g(x)) \cdot J_{\mathrm{id}\times g}(x)$$

Hier ist $J_\phi(x, g(x)) = \nabla\phi(x, g(x))$. Da ϕ an der Stelle $a = (p, q)$ ein lokales Extremum hat, hat Φ an der Stelle p ein lokales Extremum, also gilt nach dem notwendigen Extremwertkriterium 10.5.3 (ohne Nebenbedingungen!), dass $\nabla\Phi(p) = 0$.

Das Matrizenprodukt $J_\phi(p, g(p)) \cdot J_{\mathrm{id}\times g}(p)$ ist dann

$$\begin{aligned} & \frac{\partial \phi}{\partial x}(p, q) + \frac{\partial \phi}{\partial y}(p, q) \cdot \frac{\partial g}{\partial x}(p) \\ = \quad & \frac{\partial \phi}{\partial x}(p, q) - \frac{\partial \phi}{\partial y}(p, q) \cdot \left(\frac{\partial F}{\partial y}(p, q)\right)^{-1} \cdot \frac{\partial F}{\partial x}(p, q) = 0. \end{aligned}$$

Außerdem gilt (offensichtlich)

$$\frac{\partial \phi}{\partial y}(p, q) - \frac{\partial \phi}{\partial y}(p, q) \cdot \left(\frac{\partial F}{\partial y}(p, q)\right)^{-1} \cdot \frac{\partial F}{\partial y}(p, q) = 0.$$

Wir setzen dann

$$L := (\lambda_1, \dots, \lambda_m) = \frac{\partial \phi}{\partial y}(p, q) \cdot \left(\frac{\partial F}{\partial y}(p, q)\right)^{-1}$$

und erhalten aus den beiden obigen Gleichungen (für die Variablen $x_1, \dots, x_k$ und $y_1, \dots, y_k$)

$$\frac{\partial \phi}{\partial z}(p, q) - L \cdot \frac{\partial F}{\partial z}(p, q) = 0.$$

Das ist genau die Behauptung. □

Bemerkung 10.7.2. *geometrische Interpretation: das bedeutet, dass im Punkt p der Gradient von ϕ senkrecht auf $M = \{x \in U \mid g(x) = 0\}$ steht.*

KAPITEL 11

Wegintegrale und Parameterintegrale

11.1. Rektifizierbare Kurven und Bogenlänge

Wir betrachten noch einmal Kurven $\gamma\colon D \to \mathbb{R}^n$ auf einem Definitionsbereich $D \subseteq \mathbb{R}$ (meist ein Intervall). In der Physik werden beschreiben Kurven oft Bewegung von Massepunkten. Wegintegrale dienen dabei der Berechnung der Bogenlänge (also des zurückgelegten Wegs) oder der bei einer Bewegung in einem Kraftfeld nötigen Arbeit.

Definition 11.1.1. *Sei I ein Intervall. Für eine differenzierbare Kurve $\gamma\colon I \to \mathbb{R}^n$ und $t \in I$ heißt die Jacobi-Matrix $J_\gamma(t)$ von γ im Punkt $t \in I$* Tangentialvektor *an γ im Punkt t.*

Bemerkung 11.1.2. *Die physikalische Interpretation des Tangentialvektors ist als Beschreibung der Geschwindigkeit der durch die Kurve γ beschriebenen Bewegung. Die Richtung des Vektors ist dabei die Bewegungsrichtung, die Norm des Vektors gibt die Geschwindigkeit an. Oft ist auch die Notation $\dot\gamma(t)$ für den Tangentialvektor gebräuchlich.*

Definition 11.1.3. *Eine Kurve $\gamma\colon I \to \mathbb{R}^n$ heißt* regulär *oder* nicht-singulär, *wenn $\dot\gamma(t) \neq 0$ für alle $t \in I$ ist. Andernfalls heißt die Kurve* singulär.

Beispiel 11.1.4. *Die Neilesche Parabel $\gamma\colon \mathbb{R} \to \mathbb{R}^2\colon t \mapsto (t^2, t^3)$ ist singulär, der einzige singuläre Punkt ist bei $t = 0$.* □

Die lineare Annäherung regulärer Kurven durch ihre Tangentialvektoren kann benutzt werden, um den Schnittwinkel zwischen Kurven als Winkel zwischen den Tangentialvektoren am Schnittpunkt zu definieren:

Definition 11.1.5. *Für zwei reguläre Kurven $\gamma_j\colon I_j \to \mathbb{R}^n$, $j = 1, 2$ mit $\gamma_1(t_1) = \gamma_2(t_2)$ definieren wir den* Schnittwinkel α der Kurven im Schnittpunkt *$\gamma_1(t_1) = \gamma_2(t_2)$ durch*

$$\cos\alpha = \frac{\dot\gamma_1(t_1) \cdot \dot\gamma_2(t_2)}{\|\dot\gamma_1(t_1)\| \|\dot\gamma_2(t_2)\|}$$

Beispiel 11.1.6. *Eine Kurve kann sich auch selbst schneiden. Für die Kurve $\gamma\colon \mathbb{R} \to \mathbb{R}^2\colon t \mapsto (t^2 - 1, t^3 - t)$ ist $\gamma(1) = \gamma(-1) = (0, 0)$. Der Tangentialvektor für die Kurve im Punkt t ist*

$$\dot\gamma(t) = (2t, 3t^2 - 1),$$

d.h. $\dot\gamma(1) = (2, 2)$ und $\dot\gamma(-1) = (-2, 2)$. Damit gilt für den Selbst-Schnittwinkel der Kurve am Doppelpunkt $(0, 0)$

$$\cos\alpha = 0,$$

die beiden Kurvenabschnitte um $t = 1$ und $t = -1$ stehen also senkrecht aufeinander. □

Um die Bogenlänge einer Kurve $\gamma\colon [a, b] \to \mathbb{R}^n$ zu definieren, können wir versuchen, die Kurve durch einen Polygonzug zu approximieren. Wir unterteilen das

Intervall $[a,b]$ durch $a = t_0 < t_1 < \cdots < t_k = b$ und verbinden jeweils die Punkte $\gamma(t_i)$ und $\gamma(t_{i+1})$ durch eine gerade Strecke. Die Länge des Polygonzugs ist dann

$$\sum_{i=0}^{k-1} \|\gamma(t_{i+1}) - \gamma(t_i)\|.$$

Wir können dann versuchen, die Länge der Kurve zu berechnen, indem wir hier den Grenzwert der Polygonzug-Längen für immer feinere Unterteilungen und damit bessere Polygonzug-Approximationen nehmen.

DEFINITION 11.1.7. *Eine Kurve* $\gamma\colon [a,b] \to \mathbb{R}^n$ *heißt* rektifizierbar *mit Länge* L*, wenn zu jedem* $\epsilon > 0$ *ein* $\delta > 0$ *existiert, so dass für jede Unterteilung* $a = t_0 < t_1 < \cdots < t_k = b$ *mit* $\max_i |t_{i+1} - t_i| < \delta$ *gilt*

$$\left| \sum_{i=0}^{k-1} \|\gamma(t_{i+1}) - \gamma(t_i)\| - L \right| < \epsilon.$$

BEMERKUNG 11.1.8. *Nicht alle Kurven sind rektifizierbar. Amüsante Beispiele sind raumfüllende Kurven (z.B. Peano-Kurve, Hilbert-Kurve,...).*

ÜBUNGSAUFGABE 11.1.9. *Zeigen Sie, dass eine Kurve genau dann rektifizierbar ist, wenn die Längen der approximierenden Polygonzüge für alle Unterteilungen der Kurve beschränkt ist. Die Länge der Kurve ist dann das Supremum über alle solche Unterteilungen.*

SATZ 11.1.10. *Jede stetig differenzierbare Kurve* $\gamma\colon [a,b] \to \mathbb{R}^n$ *ist rektifizierbar mit Länge*

$$L = \int_a^b \|\dot\gamma(t)\| \mathrm{d}t.$$

BEWEIS. Wir können das Integral durch Riemann-Summen approximieren: für $\epsilon > 0$ existiert ein $\delta_1 > 0$, so dass für beliebige Unterteilungen $a = t_0 < t_1 < \cdots < t_k = b$ mit Feinheit $\leq \delta_1$ (d.h. $\max_i |t_{i+1} - t_i| < \delta$) gilt

$$\left| \int_a^b \|\dot\gamma(t)\| \mathrm{d}t - \sum_{i=0}^{k-1} \|\dot\gamma(t_{i+1})\| (t_{i+1} - t_i) \right| \leq \frac{\epsilon}{2}.$$

Wir wollen jetzt zeigen, dass ein $\delta \leq \delta_1$ existiert, so dass für alle Unterteilungen mit Feinheit $\leq \delta$ gilt

$$\left\| \frac{\gamma(t_{i+1}) - \gamma(t_i)}{t_{i+1} - t_i} - \dot\gamma(t_{i+1}) \right\| \leq \frac{\epsilon}{2(b-a)}.$$

Dafür können für zwei Zeitpunkte t, s (durch Vergleich von euklidischer und Maximum-Norm auf $\mathbb{R}^n$) zuerst abschätzen

$$\left\| \frac{\gamma(t) - \gamma(ts)}{t - s} - \dot\gamma(t) \right\| \leq \sqrt{n} \max_i \left| \frac{\gamma_i(t) - \gamma_i(s)}{t - s} - \dot\gamma_i(t) \right|.$$

Damit wird das Problem auf den Fall $n = 1$ zurückgeführt, in dem die Aussage aus der gleichmäßigen Stetigkeit von γ' und dem Mittelwertsatz der Differentialrechnung folgt.

Mit dieser Abschätzung erhalten wir

$$|\|\gamma(t_{i+1}) - \gamma(t_i)\| - \|\dot\gamma(t_{i+1})\|(t_{i+1} - t_i)| \leq \frac{t_{i+1} - t_i}{b - a} \frac{\epsilon}{2}$$

Damit folgt (mit Dreiecksungleichung)

$$\left| \int_a^b \|\dot\gamma(t)\| \mathrm{d}t - \sum_{i=0}^{k-1} \|\gamma(t_{i+1}) - \gamma(t_i)\| \right| \leq \epsilon.$$

□

Bemerkung 11.1.11. *Die physikalische Interpretation der Formel ist, dass bei der durch γ beschriebenen Bewegung die zurückgelegte Weglänge als Integral der Geschwindigkeit (Länge des Tangentialvektors) gegeben ist.*

Beispiel 11.1.12. *Sei $f\colon [a,b] \to \mathbb{R}$ eine stetig differenzierbare Funktion. Um die Länge des Graphen von f zu berechnen, betrachten wir die Kurve $\gamma_f\colon [a,b] \to \mathbb{R}^2\colon t \mapsto (t, f(t))$. Der Tangentialvektor dieser Kurve ist $\dot{\gamma_f}(t) = (1, f'(t))$. Die Kurve ist rektifizierbar mit Länge*

$$\int_a^b \sqrt{1 + (f'(t))^2}\mathrm{d}t.$$

□

Beispiel 11.1.13. *Eine Ellipse mit den Halbachsenlängen a und b kann durch*

$$\gamma\colon [0, 2\pi] \to \mathbb{R}^2\colon t \mapsto (a\cos t, b\sin t)$$

parametrisiert werden. Der Tangentialvektor ist dann $\dot\gamma(t) = (-a\sin t, b\cos t)$, und die Bogenlänge

$$L = \int_0^{2\pi} \sqrt{a^2\sin^2 t + b^2\cos^2 t}\mathrm{d}t = 4a\int_0^{\frac{\pi}{2}} \sqrt{1 - k^2\cos^2 t}\mathrm{d}t,$$

wobei wir im zweiten Schritt $k^2 = 1 - \left(\frac{b}{a}\right)^2$ benutzen, um a aus dem Integral herausziehen zu können. Die Zahl k heißt numerische Exzentrizität *und misst die Abweichung der Ellipse vom Kreis. Durch die Substitution $\theta = \frac{\pi}{2} - t$ (mit Differentialtransformation $\mathrm{d}t = -\mathrm{d}\theta$ geht dann die Bogenlänge in die Normalform elliptischer Integrale über*

$$L = 4a\int_0^{\frac{\pi}{2}} \sqrt{1 - k^2\sin^2\theta}\mathrm{d}\theta.$$

Diese Bogenlängeformel für die Ellipse gibt den elliptischen Integralen ihren Namen, s. auch die Diskussion des Fadenpendels in Abschnitt 8.3. □

Für eine Kurve $\gamma\colon [a,b] \to \mathbb{R}^n$ und eine bijektive stetige Abbildung $\phi\colon [c,d] \to [a,b]$ erhalten wir eine neue Kurve $\gamma\circ\phi\colon [c,d] \to \mathbb{R}^n$. Dabei wird derselbe Weg durchlaufen, aber eventuell in anderer Geschwindigkeit, das wird als *umparametrisieren* bezeichnet. Wenn ϕ und ϕ^{-1} stetig differenzierbar sind, können wir mit der Kettenregel die Auswirkungen des Umparametrisierens auf die Tangentialvektoren beschreiben:

$$J_{\gamma\circ\phi}(t) = J_\gamma(\phi(t)) \cdot \phi'(t).$$

Das bedeutet, dass sich die Tangentialvektoren von γ und $\gamma\circ\phi$ nur durch den skalaren Faktor $\phi'(t)$ unterscheiden. (Hier gibt es zwei Fälle: wenn $\phi'(t) > 0$ für alle t heißt die Umparametrisierung *orientierungserhaltend*, andernfalls *orientierungsumkehrend*. Da ϕ bijektiv und stetig differenzierbar ist, kann es keinen Vorzeichenwechsel für ϕ' geben.)

Aus der Kettenregel sehen wir auch, dass die Bogenlänge invariant unter Umparametrisierung ist. Im orientierungserhaltenden Fall ist das die Substitutionsregel für das Integral:

$$\int_a^b \|J_\gamma(t)\|\mathrm{d}t = \int_c^d \|J_{\gamma\circ\phi}(\theta)\|\mathrm{d}\theta,$$

dabei ist $\theta = \phi(t)$ und $\mathrm{d}\theta = \phi'(t)\mathrm{d}t$. Im orientierungsumkehrenden Fall hebt sich das zusätzliche Vorzeichen in $\phi'(t)$ durch Vertauschung der Integralgrenzen weg.

11.2. Vektorfelder und Wegintegral

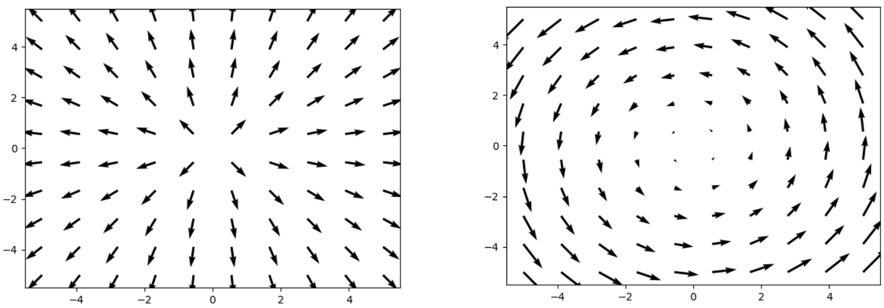
Definition 11.2.1. *Für einen offenen Definitionsbereich $D \subseteq \mathbb{R}^n$ können wir eine Abbildung $F\colon D \to \mathbb{R}^n$ so auffassen, dass jedem Punkt $x \in D$ ein Vektor $F(x) \in \mathbb{R}^n$ zugeordnet wird. Die Abbildung F heißt dann* Vektorfeld.

Wir betrachten hier typischerweise nur differenzierbare Vektorfelder.

Beispiel 11.2.2. *Vektorfelder können veranschaulicht werden, indem für ausgewählte Punkte der jeweils zugeordnete Vektor als Pfeil eingezeichnet wird:*

$$F_1(x,y) = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} \\ \frac{y}{\sqrt{x^2+^2}} \end{pmatrix}, \qquad F_2(x,y) = \begin{pmatrix} -y \\ x \end{pmatrix}.$$

□



(A) F_1 (B) F_2

Abbildung 35. Zwei Vektorfelder

Beispiel 11.2.3. *Für eine differenzierbare Funktion $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$ offen, ist $F := \nabla f\colon D \to \mathbb{R}^n$ ein Vektorfeld. Solche Vektorfelder heißen* Gradientenfelder. *Eine Funktion $U\colon D \to \mathbb{R}$ mit $F = -\nabla U$ heißt* Potential *für das Vektorfeld F.*

Ein Beispiel für ein Gradientenfeld ist Newton*sche Gravitationsfeld für eine Punktemasse m im Ursprung des $\mathbb{R}^3$. Die Gravitationskraft (auf eine Punktmasse 1 im Punkt (x,y,z)) wird in dieser Situation beschrieben durch*

$$F(x,y,z) = -\frac{Gm}{\sqrt{x^2+y^2+z^2}^3}(x,y,z).$$

Dabei ist G die Gravitationskonstante. Das zugehörige Potential ist

$$U(x,y,z) = -\frac{Gm}{\sqrt{x^2+y^2+z^2}},$$

d.h. es gilt $F = -\nabla U$. □

Definition 11.2.4. *Sei $D \subseteq \mathbb{R}^n$ offen, $F\colon D \to \mathbb{R}^n$ ein stetiges Vektorfeld und $\gamma\colon [a,b] \to D$ eine stetig differenzierbarer Kurve. Das* Wegintegral *von F entlang γ wird definiert als*

$$\int_\gamma F \cdot \mathrm{d}x := \int_a^b F(\gamma(t)) \cdot \dot{\gamma}(t) \mathrm{d}t.$$

Unter unseren Voraussetzungen kann das Wegintegral $\int_\gamma F\cdot \mathrm{d}x$ mit dem Grenzwert der Summen

$$\sum_{i=0}^{k-1} F(\gamma(\tau_i))\cdot(\gamma(t_{i+1})-\gamma(t_i))$$

über immer feinere Zerlegungen $a=t_0<t_1<\cdots<t_k=b$ und mit $\tau_i\in[t_i,t_{i+1}]$ identifiziert werden. Dieser Grenzwert existiert auch unter allgemeineren Voraussetzungen, das werden wir aber hier nicht betrachten.

Bemerkung 11.2.5. *Vektorfelder treten in der Physik zum Beispiel bei der Modellierung von Kraftfeldern (elektromagnetisches Feld oder Gravitationsfeld) auf. Die Pfeile beschreiben Richtung und Stärke der wirkenden Kraft. Das Wegintegral beschreibt die bei der Bewegung auf der Kurve γ verrichtete Arbeit.*

Flusslinien

Proposition 11.2.6. *Das Wegintegral ist linear in F, d.h.*

$$\int_\gamma (F+G)\cdot \mathrm{d}x=\int_\gamma F\cdot \mathrm{d}x+\int_\gamma G\cdot \mathrm{d}x,\qquad \int_\gamma \lambda F\cdot \mathrm{d}x=\lambda\int_\gamma F\cdot \mathrm{d}x.$$

Das Wegintegral ist additiv in γ, d.h. beim Zusammensetzen von Wegen addieren sich die Wegintegrale:

$$\int_{\gamma_1+\gamma_2} F\cdot \mathrm{d}x=\int_{\gamma_1} F\cdot \mathrm{d}x+\int_{\gamma_2} F\cdot \mathrm{d}x,\qquad \int_{-\gamma} F\cdot \mathrm{d}x=-\int_\gamma F\cdot \mathrm{d}x$$

Das Wegintegral ist beschränkt durch die maximale Norm von Vektoren entlang der Kurve γ und der Bogenlänge γ:

$$\left|\int_\gamma F\cdot \mathrm{d}x\right|\le L(\gamma)\sup_{t\in[a,b]}\|F(\gamma(t))\|$$

Beweis. Direkt aus Rechenregeln für Integrale von Funktionen in einer Variablen. □

Auch hier folgt (wie bei der Bogenlänge) aus der Kettenregel, dass stetig differenzierbares umparametrisieren das Wegintegral nicht ändert.

Satz 11.2.7. *Sei $f\colon D\to\mathbb{R}$ stetig differenzierbar, $D\subseteq\mathbb{R}^n$ offen. Wir nehmen an, dass D wegzusammenhängend ist, d.h. dass für je zwei Punkte $x,y\in D$ eine stetige Kurve $\gamma\colon[a,b]\to D$ mit $\gamma(a)=x$ und $\gamma(b)=y$ existiert. Dann ist für jede stetig differenzierbare Kurve γ, die x und y verbindet*

$$\int_\gamma \nabla f\cdot \mathrm{d}x=f(y)-f(x),$$

insbesondere ist das Wegintegral von der Kurve γ unabhängig.

Umgekehrt ist jedes Vektorfeld F auf D mit wegunabhängigen Wegintegralen ein Gradientenfeld. Für einen festgewählten Punkt $p\in D$ gilt $F=\nabla f$ mit $f(q)=\int_\gamma F\cdot \mathrm{d}x$, wobei γ eine beliebige stetig differenzierbare Kurve ist, die p mit q verbindet.

Beweis. Die Wegunabhängigkeit folgt mit einer direkten Anwendung der Kettenregel: wir betrachten die zusammengesetzte Funktion $f\circ\gamma\colon[a,b]\to\mathbb{R}$, deren Ableitung durch

$$\frac{\mathrm{d}(f\circ\gamma)}{\mathrm{d}t}=\nabla f(\gamma(t))\cdot\dot\gamma(t)$$

gegeben ist. Dann ist mit dem Hauptsatz der Differential- und Integralrechnung

$$\int_\gamma \nabla f\cdot \mathrm{d}x=\int_a^b \nabla f(\gamma(t))\cdot\dot\gamma(t)\mathrm{d}t=\int_a^b (f\circ\gamma)'(t)\mathrm{d}t=f(\gamma(b))-f(\gamma(a))=f(y)-f(x).$$

Für die zweite Aussage ist die Differenzierbarkeit von f zu zeigen, dies folgt aus Beschränktheitsabschätzungen für das Wegintegral. □

Bemerkung 11.2.8. *Es gibt noch ein anderes Kriterium, um zu prüfen, ob ein Vektorfeld ein Gradientenfeld ist (also ein Potential hat): Für ein stetig differenzierbares Gradientenfeld* $F\colon D \to \mathbb{R}^n$, $D \subseteq \mathbb{R}^n$ *offen, gilt (nach dem Satz von Schwarz)*

$$\frac{\partial F_j}{\partial x_k} = \frac{\partial F_k}{\partial x_j}$$

für alle $j, k = 1, \dots, n$. *Wenn* D *offen und konvex (oder allgemeiner offen und sternförmig) ist, folgt aus dieser Symmetrie der Ableitungen auch, dass* F *ein Gradientenfeld ist. Konzeptionelle Einordnung dieser Aussage erfolgt mit den Integralsätzen in der Analysis 3 bzw. mit dem Differentialformenkalkül in der Analysis auf Mannigfaltigkeiten.*

Beispiel 11.2.9. *Beispiele für Vektorfelder, die keine Gradientenfelder sind:*

$$F_1(x,y) = (-y, x), \qquad F_2(x,y) = (y, y-x)$$

In beiden Fällen zu sehen durch die Asymmetrie der Ableitungen. Dass F_1 *kein Gradientenfeld ist, können wir auch daran sehen, dass das Wegintegral über den Einheitskreis nicht verschwindet.* □

11.3. Integrale mit Parametern, Parameterabhängigkeit

Wir betrachten jetzt Integrale, die von einem Parameter abhängen. Das bedeutet, wir wollen für eine Funktion $f(x,y)$ in zwei Variablen - wie beim partiellen Ableiten - eine Variable y festhalten, und bezüglich x integrieren. Variieren wir dann y, erhalten wir eine Funktion

$$y \mapsto \int_a^b f(x,y)\mathrm{d}x.$$

Wir wollen verstehen, wie diese Funktion von y abhängt, insbesondere ob diese Funktion stetig bzw. differenzierbar ist.

Lemma 11.3.1. *Sei* $[a,b] \subseteq \mathbb{R}$ *ein abgeschlossenes Intervall,* $D \subseteq \mathbb{R}^m$ *beliebig, und sei* $f\colon [a,b] \times D \to \mathbb{R}$ *stetig. Sei* $(y_n)_n$ *eine Folge in* D, *die gegen einen Grenzwert* $c = \lim_{n\to\infty} y_n \in D$ *konvergiert. Dann konvergieren die Funktionen*

$$F_n\colon [a,b] \to \mathbb{R}\colon x \mapsto F_n(x) = f(x, y_n)$$

gleichmäßig gegen die Funktion $F\colon [a,b] \to \mathbb{R}\colon x \mapsto f(x,c)$.

Beweis. Die Menge $Q = \{y_n \mid n \in \mathbb{N}\} \cup \{c\}$ ist kompakt (s. Übungsaufgabe 4(a) auf Blatt 5), also nach dem Satz von Heine–Borel beschränkt und abgeschlossen. Dann ist auch $[a,b] \times Q$ beschränkt und abgeschlossen in $\mathbb{R}^{m+1}$, also nach Heine–Borel kompakt. Damit ist die Abbildung f auf $[a,b] \times Q$ gleichmäßig stetig, d.h. für alle $\epsilon > 0$ existiert ein $\delta > 0$, so dass für alle $(x,y), (x',y') \in [a,b] \times Q$ mit $\|(x,y) - (x',y')\| < \delta$ gilt $|f(x,y) - f(x',y')| < \epsilon$. Für vorgegebenes δ existiert wegen der Konvergenz $\lim_{n\to\infty} y_n = c$ ein N mit $\|c - y_n\| < \delta$ für alle $n \geq N$. Also existiert für jedes $\epsilon > 0$ ein N, so dass $|F(x) - F_n(x)| = |f(x,c) - f(x,y_n)| < \epsilon$ für alle $x \in [a,b]$ und alle $n \geq N$. Damit haben wir die gleichmäßige Konvergenz der Funktionenfolge (F_n) gezeigt. □

Satz 11.3.2 (stetige Abhängigkeit von Parameter). *Sei* $[a,b] \subseteq \mathbb{R}$ *ein abgeschlossenes Intervall,* $D \subseteq \mathbb{R}^m$ *beliebig, und sei* $f\colon [a,b] \times D \to \mathbb{R}$ *eine stetige Funktion. Dann ist die Funktion*

$$\phi\colon D \to \mathbb{R}\colon y \mapsto \int_a^b f(x,y)\mathrm{d}x$$

stetig.

Beweis. Wir zeigen Folgenstetigkeit und benutzen Proposition 9.5.13. Für eine Folge $(y_n)_n$ in D mit $\lim_{n\to\infty} y_n = c \in D$ haben wir

$$\phi(y_n) = \int_a^b f(x, y_n)\mathrm{d}x, \qquad \phi(c) = \int_a^b f(x, c)\mathrm{d}x.$$

Wegen der gleichmäßigen Konvergenz aus Lemma 11.3.1 und Proposition 7.7.3 haben wir

$$\lim_{n\to\infty} \phi(y_n) = \lim_{n\to\infty} \int_a^b f(x, y_n)\mathrm{d}x = \int_a^b \lim_{n\to\infty} f(x, y_n)\mathrm{d}x = \int_a^b f(x, c)\mathrm{d}x = \phi(c).$$

□

Lemma 11.3.3. *Seien $I, J \subseteq \mathbb{R}$ abgeschlossene Intervalle, $f\colon I \times J \to \mathbb{R}$ eine stetige Funktion, die im zweiten Argument stetig partiell differenzierbar ist. Für eine Folge $(y_n)_n$ in J mit $\lim_{n\to\infty} y_n = c$ und $y_n \neq c$ für alle n definieren wir die folgenden Funktionen*

$$F_n\colon I \to \mathbb{R}\colon x \mapsto \frac{f(x, y_n) - f(x, c)}{y_n - c}, \qquad F\colon I \to \mathbb{R}\colon x \mapsto \partial_y f(x, c).$$

Dann konvergiert die Funktionenfolge (F_n) auf I gleichmäßig gegen F.

Beweis. Da $I \times J$ kompakt ist, ist $\partial_y f\colon I \times J \to \mathbb{R}$ sogar gleichmäßig stetig, d.h. für jedes $\epsilon > 0$ existiert ein $\delta > 0$, so dass für alle $y, y' \in J$ mit $|y - y'| < \delta$ gilt $|\partial_y f(x, y) - \partial_y f(x, y')| < \epsilon$. Für die Differenzenquotienten gibt es nach dem Mittelwertsatz der Differentialrechnung 6.3.4 für jedes $x \in I$ und $n \in \mathbb{N}$ ein $\xi_n \in [c, y_n]$ mit $F_n(x) = \partial_y f(x, \xi_n)$. Wegen $\lim_{n\to\infty} y_n = c$ existiert für vorgegebenes $\delta > 0$ ein N mit $|c - y_n| < \delta$ für alle $n \geq N$, damit auch $|c - \xi_n| < \delta$ für alle $n \geq N$. Aus der gleichmäßigen Stetigkeit folgt dann die gleichmäßige Konvergenz:

$$|F(x) - F_n(x)| = |\partial_y f(x, c) - \partial_y f(x, \xi_n)| < \epsilon$$

für alle $x \in I$ und $n \geq N$. □

Satz 11.3.4 (differenzierbare Abhängigkeit vom Parameter). *Seien $I, J \subseteq \mathbb{R}$ abgeschlossene Intervalle und $f\colon I \times J \to \mathbb{R}$ eine stetige Funktion, die im zweiten Argument stetig partiell differenzierbar ist. Dann ist die Funktion*

$$\phi\colon J \to \mathbb{R}\colon y \mapsto \int_I f(x, y)\mathrm{d}x$$

stetig differenzierbar und es gilt

$$\frac{\mathrm{d}\phi(y)}{\mathrm{d}y} = \int_I \partial_y f(x, y)\mathrm{d}x.$$

Beweis. Aus Lemma 11.3.3 und Proposition 7.7.3 folgt

$$\phi'(c) = \lim_{n\to\infty} \frac{\phi(y_n) - \phi(c)}{y_n - c} = \lim_{n\to\infty} \int_I \frac{f(x, y_n) - f(x, c)}{y_n - c}\mathrm{d}x = \int_I \partial_y f(x, c)\mathrm{d}x.$$

Die Ableitung $\phi'(c)$ existiert also für alle $c \in I$, hat die behauptete Form und ist nach Satz 11.3.2 stetig. □

Bemerkung 11.3.5. *Das funktioniert allgemeiner auch für Funktionen in mehreren Veränderlichen: für ein abgeschlossenes Intervall $[a, b] \subseteq \mathbb{R}$, eine offene Teilmenge $D \subseteq \mathbb{R}^n$ und eine stetig partiell differenzierbare Funktion $f\colon [a, b] \times D \to \mathbb{R}$ ist*

$$\partial_{y_i} \int_a^b f(x, y_1, \ldots, y_n)\mathrm{d}x = \int_a^b \partial_{y_i} f(x, y_1, \ldots, y_n)\mathrm{d}x.$$

Beispiel 11.3.6 (Potentialbestimmung durch Parameterintegrale). *Sei* $D = \{x \in \mathbb{R}^n \mid \|x\| < r\}$ *ein offener Ball von Radius* r *um den Ursprung, und* $F\colon D \to \mathbb{R}^n$ *ein stetig differenzierbares Vektorfeld mit*

$$\frac{\partial F_j}{\partial x_k} = \frac{\partial F_k}{\partial x_j}$$

für alle $j, k = 1, \ldots, n$. *Dann ist die Funktion*

$$U(x) = -\sum_{i=1}^{n} \left(\int_0^1 F_i(tx)\mathrm{d}t \right) x_i$$

ein Potential für F, *d.h.* $F = -\nabla U$. *Nach Satz 11.3.4 haben wir*

$$\begin{aligned}
\partial_{x_j} f(x) &= \sum_{i=1}^{n} \left(\partial_{x_j} \int_0^1 F_i(tx)\mathrm{d}t \right) x_i + \sum_{i=1}^{n} \left(\int_0^1 F_i(tx)\mathrm{d}t \right) \partial_{x_j} x_i \\
&= \sum_{i=1}^{n} \left(\int_0^1 t\partial_{x_j} F_i(tx)\mathrm{d}t \right) x_i + \int_0^1 F_j(tx)\mathrm{d}t \\
&= \int_0^1 \left(t \sum_{i=1}^{n} \partial_{x_j} F_i(tx) x_i + F_j(tx) \right) \mathrm{d}t.
\end{aligned}$$

Für festes $x \in D$ *ist*

$$\begin{aligned}
\frac{\mathrm{d}(tF_j(tx))}{\mathrm{d}t} &= F_j(tx) + t\frac{\mathrm{d}(F_j(tx))}{\mathrm{d}t} \\
&= F_j(tx) + t\sum_{i=1}^{n} \partial_i F_j(tx) x_i = F_j(tx) + t\sum_{i=1}^{n} \partial_j F_i(tx) x_i
\end{aligned}$$

Durch Kombination beider Rechnungen erhalten wir

$$\partial_{x_j} f(x) = \int_0^1 \frac{\mathrm{d}(tF_j(tx))}{\mathrm{d}t} \mathrm{d}t = tF_j(tx)\Big|_0^1 = F_j(x).$$

□

11.4. Doppelintegrale

Durch die Stetigkeitsaussage von Satz 11.3.2 können wir nun auch Funktionen $f(x_1, \ldots, x_n)$ bezüglich verschiedener Variablen integrieren. Betrachten wir erst den Fall von zwei Variablen, d.h. eine stetige Funktion $f\colon [a, b] \times [c, d] \to \mathbb{R}$. Nach Satz 11.3.2 ist

$$F(y) = \int_a^b f(x, y)\mathrm{d}x$$

stetig. Dann kann die Funktion auch noch bezüglich y integriert werden, und wir erhalten ein Doppelintegral

$$\iint_{[a,b]\times[c,d]} f(x, y)\mathrm{d}x\mathrm{d}y := \int_c^d \left(\int_a^b f(x, y)\mathrm{d}x \right) \mathrm{d}y.$$

Die Integrationsreihenfolge kann hier vertauscht werden. Eine viel allgemeinere Aussage dazu wird in der Analysis 3 gemacht (Satz von Fubini).

Satz 11.4.1. *Sei* $f\colon [a, b] \times [c, d] \to \mathbb{R}$ *stetig. Dann gilt*

$$\int_c^d \left(\int_a^b f(x, y)\mathrm{d}x \right) \mathrm{d}y = \int_a^b \left(\int_c^d f(x, y)\mathrm{d}y \right) \mathrm{d}x$$

Beweis. Hier wird die Differenzierbarkeit von Parameterintegralen genutzt. Wir definieren die Hilfsfunktion

$$\phi(y) = \int_a^b \left(\int_c^y f(x,t)\mathrm{d}t \right) \mathrm{d}x$$

Nach Satz 11.3.4 ist die Funktion ϕ differenzierbar und

$$\phi'(y) = \int_a^b \partial_y \left(\int_c^y f(x,t)\mathrm{d}t \right) \mathrm{d}x = \int_a^b f(x,y)\mathrm{d}x.$$

Dann gilt (mit $\phi(c) = 0$)

$$\int_c^d \left(\int_a^b f(x,y)\mathrm{d}x \right) \mathrm{d}y = \int_c^d \phi'(y)\mathrm{d}y = \phi(d) = \int_a^b \left(\int_c^d f(x,y)\mathrm{d}y \right) \mathrm{d}x.$$

□

Bemerkung 11.4.2. *Allgemeiner können wir Mehrfachintegrale für Funktionen*

$$f\colon [a_1, b_1] \times \cdots \times [a_n, b_n] \to \mathbb{R}$$

definieren:

$$\int_{a_n}^{b_n} \cdots \int_{a_1}^{b_1} f(x_1, \ldots, x_n)\mathrm{d}x_1 \cdots \mathrm{d}x_n.$$

Eine Modifikation ergibt sich, wenn für das innere Integral auch variable Grenzen zugelassen werden:

Definition 11.4.3 (Integration über Normalbereiche). *Eine Teilmenge $B_1 \subseteq \mathbb{R}^2$ heißt* Normalbereich erster Art*, wenn es $a, b \in \mathbb{R}$ und stetig differenzierbare Funktionen $g, h\colon [a,b] \to \mathbb{R}$ gibt, so dass $g(x) \leq h(x)$ für alle $x \in [a,b]$ und*

$$B_1 = \{(x,y) \in \mathbb{R}^2 \mid a \leq x \leq b, g(x) \leq y \leq h(x)\}.$$

Für eine auf B_1 stetige Funktion f definieren wir dann

$$\iint_{B_1} f(x,y)\mathrm{d}x\mathrm{d}y = \int_a^b \left(\int_{g(x)}^{h(x)} f(x,y)\mathrm{d}y \right) \mathrm{d}x.$$

Eine Teilmenge $B_2 \subseteq \mathbb{R}^2$ heißt Normalbereich zweiter Art*, wenn es $c, d \in \mathbb{R}$ und stetig differenzierbare Funktionen $r, l\colon [c,d] \to \mathbb{R}$ gibt, so dass $l(y) \leq r(y)$ für alle $y \in [c,d]$ und*

$$B_2 = \{(x,y) \in \mathbb{R}^2 \mid l(y) \leq x \leq r(y), c \leq y \leq d\}.$$

Für eine auf B_2 stetige Funktion f definieren wir dann

$$\iint_{B_2} f(x,y)\mathrm{d}x\mathrm{d}y = \int_c^d \left(\int_{l(y)}^{r(y)} f(x,y)\mathrm{d}x \right) \mathrm{d}y.$$

Solche Integrale können zur Berechnung von (speziellen) Flächen und Volumina benutzt werden. Die allgemeine Integrationstheorie wird in der Analysis 3 aufgebaut.

Beispiel 11.4.4. *Wir wollen die Fläche B berechnen, die in $\mathbb{R}^2$ von den Graphen der Funktionen $h(x) = 4 - x^2$ und $g(x) = x + 2$ für $x \in [-2, 1]$ eingeschlossen* *wird. Dies ist ein Normalbereich erster Art. Der Flächeninhalt/das Doppelintegral ist*

$$\begin{aligned}\iint_B \mathrm{d}x\mathrm{d}y &= \int_{-2}^{1}\left(\int_{x+2}^{4-x^2}\mathrm{d}y\right)\mathrm{d}x \\ &= \int_{-2}^{1}\left(y|_{x+2}^{4-x^2}\right)\mathrm{d}x \\ &= \int_{-2}^{1}\left(4-x^2-x-2\right)\mathrm{d}x = \left[2x-\frac{x^3}{3}-\frac{x^2}{2}\right]_{-2}^{1} = \frac{9}{2}.\end{aligned}$$

Alternativ können wir auch eine Funktion, z.B. $f(x) = x$, über diesen Normalbereich integrieren:

$$\begin{aligned}\iint_B f(x)\mathrm{d}x\mathrm{d}y &= \int_{-2}^{1}\left(\int_{x+2}^{4-x^2}x\mathrm{d}y\right)\mathrm{d}x \\ &= \int_{-2}^{1}\left(xy|_{x+2}^{4-x^2}\right)\mathrm{d}x \\ &= \int_{-2}^{1}\left(4x-x^3-x^2-2x\right)\mathrm{d}x = \left[x^2-\frac{x^3}{3}-\frac{x^4}{4}\right]_{-2}^{1} = -\frac{9}{4}.\end{aligned}$$

□

BEISPIEL 11.4.5. *Für einen Normalbereich B mit einer Dichtefunktion $\mu\colon B \to \mathbb{R}$ berechnet das Integral $\iint_B \mu(x,y)\mathrm{d}x\mathrm{d}y$ die in B verteilte Masse M von B. Der Schwerpunkt dieses zwei-dimensionalen Objekts ist $S = (x_S, y_S)$ mit den Koordinaten*

$$x_s = \frac{1}{M}\iint_B x\mu(x,y)\mathrm{d}x\mathrm{d}y, \qquad y_s = \frac{1}{M}\iint_B y\mu(x,y)\mathrm{d}x\mathrm{d}y$$

Interpretieren wir das vorangegangene Beispiel in diesem Kontext als Bereich B mit der Dichtefunktion $\mu(x,y) = 1$, ist die Masse M gleich dem Flächeninhalt von B, d.h. $M = \frac{9}{2}$. Die x-Koordinate des Schwerpunkts dann

$$x_S = \frac{1}{M}\iint_B x\mathrm{d}x\mathrm{d}y = \frac{2}{9}\left(-\frac{9}{4}\right) = \frac{1}{2}.$$

Mit einer ähnlichen Rechnung ist die y-Koordinate des Schwerpunkts

$$y_S = \frac{1}{M}\iint_B y\mathrm{d}x\mathrm{d}y = \frac{2}{9}\cdot\frac{54}{5} = \frac{12}{5}.$$

□

BEISPIEL 11.4.6. *Ähnlich können auch Mehrfachintegrale berechnet werden. Ein Beispiel, das bei der Berechnung des Kugelvolumens in Polarkoordinaten auftritt ist:*

$$\begin{aligned}\int_0^R\int_0^\pi\int_0^{2\pi} r^2\sin\theta\mathrm{d}\phi\mathrm{d}\theta\mathrm{d}r &= 2\pi\int_0^R r^2\left(\int_0^\pi \sin\theta\mathrm{d}\theta\right)\mathrm{d}r \\ &= 2\pi\frac{R^3}{3}[-\cos\theta]_0^\pi = \frac{4\pi}{3}R^3.\end{aligned}$$

□

ANHANG A

Grundlagen lineare Algebra

A.1. Vektorräume

DEFINITION A.1.1. *Sei K ein Körper. Ein K-*Vektorraum *ist eine Menge V zusammen mit zwei Operationen*

Addition: $+\colon V \times V \to V\colon (a,b) \mapsto a+b$ *und*
Skalarmultiplikation: $\cdot\colon K \times V \to V\colon (\lambda, v) \mapsto \lambda \cdot v$,

so dass die folgenden Aussagen gelten:

(1) *$(V,+)$ ist eine abelsche Gruppe. Das neutrale Element wird mit 0 bezeichnet, das inverse Element zu v mit $-v$.*
(2) *Für $\lambda, \mu \in K$ und $v, w \in V$ gelten*

$$\begin{aligned} (\lambda+\mu)\cdot v &= \lambda\cdot v + \mu \cdot v \\ \lambda\cdot(v+w) &= \lambda\cdot v + \lambda\cdot w \\ \lambda\cdot(\mu\cdot v) &= (\lambda\mu)\cdot v \\ 1\cdot v &= v \end{aligned}$$

BEISPIEL A.1.2.

- *Der Vektorraum*
$$K^n = \underbrace{K \times \cdots \times K}_{n \text{ Faktoren}} = \{(x_1,\ldots,x_n) \mid x_i \in K\}$$
besteht aus den geordneten n-Tupeln von Elementen aus K. Addition und Skalarmultiplikation erfolgen komponentenweise:
$$\lambda\cdot(x_1,\ldots,x_n) + \mu\cdot(y_1,\ldots,y_n) = (\lambda x_1 + \mu y_1, \ldots, \lambda x_n + \mu y_n).$$
- *$\mathbb{C}$ ist ein Vektorraum über $\mathbb{R}$, mit der Addition komplexer Zahlen und Skalarmultiplikation durch reelle Zahlen $\lambda\cdot(x+\mathrm{i}y) = \lambda x + \mathrm{i}\lambda y$.*
- *Die Menge der reellen Folgen $(a_n)_n$ ist ein $\mathbb{R}$-Vektorraum, mit der Addition $(a_n)_n + (b_n)_n = (a_n+b_n)_n$ und Skalarmultiplikation $\lambda\cdot(a_n)_n = (\lambda a_n)_n$.*
- *Sei $I \subseteq \mathbb{R}$ ein Intervall. Die Menge $C^n(I)$ der n-mal stetig differenzierbaren Funktionen auf I ist ein $\mathbb{R}$-Vektorraum. Addition und Skalarmultiplikation sind dabei durch*
$$(f+g)(x) = f(x) + g(x) \text{ und } (\lambda\cdot f)(x) = \lambda f(x)$$
gegeben.
- *Für ein abgeschlossenes Intervall $I \subseteq \mathbb{R}$ ist die Menge $\mathcal{T}(I)$ der reellwertigen Treppenfunktionen auf I ein $\mathbb{R}$-Vektorraum. Ebenso ist die Menge $\mathcal{R}(I)$ der Regelfunktionen auf I ein $\mathbb{R}$-Vektorraum.*

□

DEFINITION A.1.3. *Sei K ein Körper und V ein K-Vektorraum. Eine Teilmenge $W \subseteq V$ von K heißt* Untervektorraum *von V, wenn sie unter Addition und Skalarmultiplikation abgeschlossen ist.*

BEISPIEL A.1.4.

- *Sei $I \subseteq \mathbb{R}$ ein Intervall. Für natürliche Zahlen $n > m$ ist $C^n(I) \subset C^m(I)$ ein Untervektorraum.*

- *Die Menge der konvergenten Folgen ist ein Untervektorraum im Vektorraum aller Folgen. Im Vektorraum der konvergenten Folgen bilden die Nullfolgen einen Untervektorraum.*
- *Für ein abgeschlossenes Intervall $I \subseteq \mathbb{R}$ ist die Menge der Treppenfunktionen $\mathcal{T}(I)$ ein Untervektorraum im Vektorraum $\mathcal{R}(I)$ der Regelfunktionen.*

□

ÜBUNGSAUFGABE A.1.5. *Im Vektorraum der Nullfolgen bilden die quadratsummierbaren Folgen, d.h. die Folgen $(a_n)_n$ für die $\sum_{n=0}^{\infty} a_n^2$ konvergiert, einen Untervektorraum.*

Bilden die Folgen mit Grenzwert 1 einen Untervektorraum des Raums der Folgen?

lineare Abhängigkeit, Erzeugendensystem, Basis
Existenz einer Basis, Dimension, Basisergänzungssatz

A.2. Lineare Abbildungen

DEFINITION A.2.1. *Sei K ein Körper, und seien V, W zwei K-Vektorräume. Eine Abbildung $f\colon V \to W$ heißt* lineare Abbildung, *wenn für alle $\lambda, \mu \in K$ und $v, w \in V$ gilt*

$$f(\lambda v + \mu w) = \lambda f(v) + \mu f(w).$$

BEMERKUNG A.2.2. *Das bedeutet, dass eine Abbildung $f\colon V \to W$ linear ist, wenn sie mit Addition und Skalarmultiplikation kompatibel ist.*

BEISPIEL A.2.3.

- *Bezeichnen wir mit $\mathcal{F}$ den Vektorraum der konvergenten Folgen, dann ist "Grenzwert bilden" eine lineare Abbildung*

$$\lim_{n\to\infty}(-)\colon \mathcal{F} \to \mathbb{R}\colon (a_n)_n \mapsto \lim_{n\to\infty} a_n$$

- *Sei I ein Intervall. Für jedes $c \in I$ ist "Auswerten an c" eine lineare Abbildung*

$$C^0(I) \to \mathbb{R}\colon f \mapsto f(c).$$

auf dem Vektorraum $C^0(I)$ der reell-wertigen stetigen Funktionen $f\colon I \to \mathbb{R}$.

- *Sei I ein Intervall. Für jede natürliche Zahl liefert die Ableitung eine lineare Abbildung*

$$C^n(I) \to C^{n-1}(I)\colon f \mapsto f'$$

- *Für ein abgeschlossenes Intervall $[a, b]$ ist das Integral auf dem Raum $\mathcal{R}([a, b])$ der Regelfunktionen eine lineare Abbildung*

$$\mathcal{R}([a,b]) \to \mathbb{R}\colon (f\colon [a,b] \to \mathbb{R}) \mapsto \int_a^b f(x)\mathrm{d}x.$$

□

ÜBUNGSAUFGABE A.2.4. *Welche der folgenden Abbildungen sind linear?*

(1) Ableitung von Polynomen $\partial : \mathbb{R}[X] \to \mathbb{R}[X]$
(2) $\mathbb{R}^1 \to \mathbb{R}^1 : x \mapsto x^2$
(3) Skalarprodukt

$$\langle v, - \rangle : \mathbb{R}^3 \to \mathbb{R} : w \mapsto \langle v, w \rangle$$

für gegebenen Vektor $v \in \mathbb{R}^3$.

darstellende Matrizen

ÜBUNGSAUFGABE A.2.5. *(1) Geben Sie die darstellende Matrix für die Spiegelung an der Ebene $x_1 + x_2 + x_3 = 0$ bezüglich der Standardbasis im $\mathbb{R}^3$ an.*

(2) Bezeichne $\mathbb{R}[X]_{\leq 4}$ den Vektorraum der reellen Polynome vom Grad ≤ 4. Geben Sie die darstellende Matrix der Ableitungsabbildung $\partial : \mathbb{R}[X]_{\leq 4} \to \mathbb{R}[X]_{\leq 4}$ bezüglich der Basis $\{1, 1+X, 1+X+X^2, \ldots, 1+X+X^2+X^3+X^4\}$ an.

(3) Wir betrachten $\mathbb{C}$ als $\mathbb{R}$-Vektorraum mit der Basis $\{1, \mathrm{i}\}$. Geben Sie die darstellende Matrix für die Multiplikation mit der komplexen Zahl $z = a + b\mathrm{i}$ an.

Isomorphismus

Kern, Bild, Dimensionsformel, Rang einer linearen Abbildung

Determinante, Definition, Multiplikativität, Laplace-Entwicklungssatz

Eigenwerte, Eigenvektoren, symmetrische Matrizen haben reelle Eigenwerte, Diagonalisierbarkeit selbstadjungierter Endomorphismen, symmetrische Matrizen

Linearformen, Bilinearformen

Skalarprodukte, Orthonormalbasen, Gram–Schmidt

orthogonales Komplement, Zerlegung von Vektoren, Projektion auf Unterraum